stochastic processes and their applications

ELSEVIER Stochastic Processes and their Applications 58 (1995) 329–360

Remarques sur l'ergodicité des algorithmes de recuit simulé sur un graphe

Laurent Miclo [1]

Laboratoire de Statistique et Probabilités, Université Paul Sabatier et CNRS, 118, route de Narbonne, 31062 Toulouse cedex, France

Received 24 January 1993; revised 22 January 1995

Abstract

We consider the simulated annealing algorithm associated to a potential U on a graph (M,q) (reversible or satisfying the Hajek's weak reversibility condition), whose temperature at time $t \geqslant 0$ is given by $k\ln^{-1}(1+t)$, with $k > c(M,U)$ the critical constant for the ergodicity in law of the process. Let $\widetilde{M}$ (respectively $\widehat{M}$) the connected component of the set $\{x \in M \,|U(x) < \min_M U + k\}$ (respectively $\{x \in M \,|U(x) \leqslant \min_M U + k\}$) which contains all the global minima. We will see that $\widehat{M}$ is the recurrent set and that the occupation times of points in $\widetilde{M}$ (or of points x_0 in $\widehat{M}$ such that $U(x_0) = k$) satisfy a strong law of large numbers. Furthermore, if the graph is a reversible tree and if $\widehat{M} = \widetilde{M}$, we shall study the behaviour in law and a.s. of the fluctuations around these laws of large numbers (central limit theorem and law of the iterated logarithm).

Résumé : On considère l'algorithme de recuit simulé associé à un potentiel U sur un graphe (M,q) (réversible ou satisfaisant la condition de réversibilité faible de Hajek), dont la décroissance de la température est en $k\ln^{-1}(1+t)$, avec $k > c(M,U)$ la constante critique assurant l'ergodicité en loi de ce processus. Soit $\widetilde{M}$ (respectivement $\widehat{M}$) la composante connexe de l'ensemble $\{x \in M \mid U(x) < \min_M U+k\}$ (respectivement $\{x \in M \mid U(x) \leqslant \min_M U+k\}$) qui contient l'ensemble des minima globaux du potentiel U. On verra que $\widehat{M}$ est l'ensemble des points récurrents et que les temps d'occupation des points de $\widetilde{M}$ (ou des points $x_0 \in \widehat{M}$ tels que $U(x_0) = k$) satisfont une loi forte des grands nombres. Dans le cas où le graphe est un arbre réversible et où $\widehat{M} = \widetilde{M}$, on étudiera également le comportement en loi et p.s. des fluctuations autour de ces lois des grands nombres (théoréme de la limite centrale et loi du logarithme itéré).

Keywords: Simulated annealing, Law of large numbers, Central limit theorem, Law of the iterated logarithm.

1. Introduction

Soit (M,q,μ) un graphe fini irréductible et réversible: Mest un ensemble fini, μ est une probabilité chargeant tous les points de M et $q : M \times M \to \mathbb{R}_+$ est un noyau de

[1] Ce travail a été effectué à l'Institut de Recherche Mathématique Avancée, Université Louis Pasteur et CNRS, Strasbourg, France.

0304-4149/95/$09.50 © 1995 Elsevier Science B.V. All rights reserved
SSDI 0304-4149(95)00022-4

probabilités de transitions irréductible (pour tous $x, y \in M$, il existe un chemin allant de x à y, c'est-à-dire une suite $(p(i))_{1 \leqslant i \leqslant n}$ d'éléments de M telle que $p(1) = x$, $p(n) = y$ et satisfaisant pour tout $1 \leqslant i < n$, $q(p(i), p(i+1)) > 0$) et réversible par rapport à μ (i.e. pour tous $x, y \in M$, $\mu(x)q(x,y) = \mu(y)q(y,x)$). Soit également U une fonction sur M telle que $\min_{z \in M} U(z) = 0$. On s'intéresse aux algorithmes de recuit simulé associés, c'est-à-dire aux processus de Markov (càdlàg.) à valeurs dans M, $(X_t)_{t \geqslant 0}$, dont la famille des générateurs $(L_{\beta_t})_{t \geqslant 0}$ est donnée par

$$\forall\, t \geqslant 0,\ \forall\, \phi : M \to \mathbb{R},\ \forall\, x \in M,$$
$$L_{\beta_t}\phi(x) = \sum_{y \in M} (\phi(y) - \phi(x)) q_{\beta_t}(x,y),$$

où pour tout $\beta \geqslant 0$ (qui représente l'inverse de la température), le noyau de probabilités (ou plus rigoureusement ici, d'intensités) de transitions q_β est défini par

$$\forall\, x, y \in M,$$
$$q_\beta(x,y) = \begin{cases} \exp(-\beta(U(y) - U(x))_+)q(x,y) & \text{si } y \neq x, \\ 1 - \sum_{z \neq x} q_\beta(x,z) & \text{si } y = x \end{cases}$$

et où $\mathbb{R}_+ \ni t \mapsto \beta_t \in \mathbb{R}_+$ est une fonction de classe C^1 telle que $\lim_{t \to +\infty} \beta_t = 0$.

En fait, on ne considèrera ici que des évolutions de la forme

$$\beta_t = k^{-1} \ln(1+t)$$

avec $k > 0$ assez grand.

Plus précisément, pour $x, y \in M$, soit $\mathscr{C}_{x,y}$ l'ensemble des chemins allant de x à y sans se recouper (i.e. des suites $p = (p(i))_{1 \leqslant i \leqslant n}$ d'éléments distincts vérifiant les conditions énoncées un peu plus haut). On note $e(p)$ l'élévation pour U d'un tel chemin p,

$$e(p) = \max_{1 \leqslant i \leqslant n} U(p(i)),$$

puis on définit

$$H(x,y) = \min_{p \in \mathscr{C}_{x,y}} e(p)$$

et

$$c(M,U) = \max_{x,y \in M} [H(x,y) - U(x) - U(y)] \geqslant 0.$$

On supposera désormais que $k > c(M,U)$. Notons $N = \{z \in M \mid U(z) = 0\}$ l'ensemble des minima globaux de U, $\widetilde{M}$ la composante connexe (pour le noyau de probabilités de transitions q) de l'ensemble $\{x \in M \mid U(x) < k\}$ qui contient N (l'hypothèse $k > c(M,U)$ assure que N est bien inclus dans l'une des composantes connexes de cet ensemble) et $\widehat{M}$ la composante connexe de l'ensemble $\{x \in M \mid U(x) \leqslant k\}$ qui contient N. En reprenant une terminologie introduite par Freidlin et Wentzell (voir (Freidlin et Wentzell, 1984), mais surtout (Catoni, 1991; et Trouvé, 1993), l'ensemble $\widehat{M}$ (respectivement $\widetilde{M}$) est le cycle maximal (pour l'inclusion ensembliste) dont le fond est N et dont la hauteur de mélange est inférieure (respectivement strictement inférieure) à k.

On a vu dans (Miclo, 1995) que les points de $M|\widehat{M}$ étaient transcients p.s. (c'est-à-dire que p.s., $\lim_{t\to+\infty} 1_{\widehat{M}}(X_t) = 1$) et que les points de $\widetilde{M}$ étaient récurrents (i.e., $\forall\, x_0 \in \widetilde{M}$, $\limsup_{t\to+\infty} 1_{\{x_0\}}(X_t) = 1$). On va retrouver ces résultats par une autre méthode et on verra plus précisément que tous les points de $\widehat{M}$ sont récurrents.

On peut donc se demander si à l'intérieur de $\widehat{M}$, le processus $(X_t)_{t\geqslant 0}$ satisfait les propriétés ergodiques usuelles:

Soit $x_0 \in \widehat{M}$ fixé, on note ϕ l'indicatrice de $\{x_0\}$ et pour $t\geqslant 0$,

$$G_t(\phi) = \int_0^t \phi(X_s)\,\mathrm{d}s$$

le temps passé en x_0.

On cherche à trouver un équivalent de $G_t(\phi)$ en temps grand (Théorème ergodique, encore appelé loi faible ou forte des grands nombres, suivant que la limite est en probabilité ou p.s.) et le comportement asymptotique des fluctuations autour de cet équivalent (Théorème de la limite centrale et du logarithme itéré).

Malheureusement, nous n'avons pu établir un théorème ergodique que pour $x_0 \in \widetilde{M}$ ou $x_0 \in \widehat{M}$ avec $U(x_0) = k$, la question reste ouverte de comprendre le comportement de $G_t(\phi)$ pour $x_0 \in \widehat{M}|\widetilde{M}$ tel que $U(x_0) < k$. Plus grave, pour obtenir les estimées nécessaires à la démonstration de résultats sur le comportement asymptotique des fluctuations, nous avons dû nous restreindre au cas où (M, q, μ) est un arbre, situation dans laquelle tous les calculs peuvent se faire explicitement (mais qui pratiquement, est d'un intérêt réduit, voir nul), et faire l'hypothèse que $\widehat{M} = \widetilde{M}$. Cependant, du-moins sous cette dernière condition, nous pensons que les résultats présentés sur ce sujet doivent encore être vérifiés dans le cas d'un graphe irréductible et réversible quelconque.

Un objet important dans l'étude du recuit simulé, est la probabilité de Gibbs μ_{β_s} associée à la température β_s^{-1} et à la fonction U, et qui est définie par

$$\forall\, x \in M, \qquad \mu_{\beta_s}(x) = Z_{\beta_s}^{-1} \exp(-\beta_s U(x))\mu(x)$$

avec $Z_{\beta_s} = \sum_{z\in M} \exp(-\beta_s U(z))\mu(z)$. On pose également

$$g_t(\phi) = \int_0^t \mu_{\beta_s}(\phi)\,\mathrm{d}s.$$

Notons que $g_t(\phi)$ diverge en temps grand, du fait que $U(x_0)\leqslant k$, plus précisément, si $U(x_0) < k$, alors

$$g_t(\phi) \sim \mu(N)^{-1}\mu(x_0)(1 - k^{-1}U(x_0))^{-1} t^{1-k^{-1}U(x_0)}$$

et si $U(x_0) = k$,

$$g_t(\phi) \sim \mu(N)^{-1}\mu(x_0)\ln(1 + t).$$

Notons $\widehat{M} = \widetilde{M} \cup \{x \in \widehat{M} \mid U(x) = k\}$. Alors,

Théorème 1. *Pour tout* $x_0 \in \widehat{M}$, *on a p.s.*,

$$\lim_{t \to +\infty} g_t^{-1}(\phi) G_t(\phi) = 1.$$

Supposons maintenant que $\widehat{M} = \widetilde{M}$, et que (M, q, μ) soit un arbre, c'est-à-dire que pour tout $x, y \in M$, il existe un unique chemin (on sous-entend désormais, ne se recoupant pas) allant de x à y. On le notera $p_{x,y} = (p_{x,y}(i))_{1 \leqslant i \leqslant n_{x,y}}$ et $e^{\widetilde{M}}(p_{x,y})$ désignera la quantité

$$\max_{1 \leqslant i \leqslant n_{x,y} \setminus p_{x,y}(i) \in \widetilde{M}} U(p_{x,y}(i)).$$

Pour tout $x_0 \in \widetilde{M}$ fixé, définissons également

$$\widetilde{W}(x_0) = \max_{x \neq x_0} [e^{\widetilde{M}}(p_{x_0,x}) - 2U(x)] - 2U(x_0),$$

$$\underline{W}(x_0) = -\min_{x \notin \widetilde{M}} U(x) - U(x_0)$$

(par convention $\underline{W}(x_0) = -\infty$ si $\widetilde{M} = M$, i.e. si $k > \max_M U$)

$$W(x_0) = \frac{1 + k^{-1} \widetilde{W}(x_0)}{2} \vee (1 + k^{-1} \underline{W}(x_0))$$

(qui dépendent aussi de k) et remarquons que $\widetilde{W}(x_0) > -k$ dès que l'on suppose par exemple que $\widetilde{M}$ n'est pas réduit à un point. Faisons l'hypothèse que $W(x_0) > 0$ et notons pour $t > 0$,

$$F_t = t^{-W(x_0)}(G_t(\phi) - g_t(\phi))$$

les fluctuations renormalisées. Ce sont les positions relatives de $\widetilde{W}(x_0)$ et de $\underline{W}(x_0)$ qui vont déterminer leur comportement asymptotique en loi.

Théorème 2. *Sous les restrictions précédentes, il existe des nombres* $\underline{m}(x_0, k) > 0$ *et* $\widetilde{\sigma}(x_0, k) > 0$, *que l'on explicitera ultérieurement, tels que l'on ait la convergence en loi des* $\widetilde{F}_t$ *en temps grand vers une gaussienne de moyenne* $m(x_0, k)$ *et de variance* $\sigma(x_0, k)$, *avec*

$$m(x_0, k) = \begin{cases} 0 & \text{si } \widetilde{W}(x_0) > k + 2\underline{W}(x_0), \\ \underline{m}(x_0, k) & \text{sinon,} \end{cases}$$

$$\sigma(x_0, k) = \begin{cases} \widetilde{\sigma}(x_0, k) & \text{si } \widetilde{W}(x_0) \geqslant k + 2\underline{W}(x_0), \\ 0 & \text{sinon.} \end{cases}$$

Dans le cas où $\widetilde{W}(x_0) < k + 2\underline{W}(x_0)$, *on a en fait convergence p.s., voir le théorème suivant.*

On prouvera en fait une version fonctionnelle plus forte de ce résultat (convergence vers un processus gaussien (dégénéré si $\widetilde{W}(x_0) < k + 2\underline{W}(x_0)$), voir la Section 5).

De même Indiquons qu'il existe plusieurs possibilités pour le comportement asymptotique p.s. des fluctuations:

Théorème 3. *Sous les hypothèses précédentes (dont $W(x_0) > 0$),*
- *Si $\widetilde{W}(x_0) \geqslant k + 2\underline{W}(x_0)$, on a*

$$\limsup_{t \to +\infty} \frac{F_t}{\sqrt{\ln_2(t)}} = \sqrt{2\widetilde{\sigma}(x_0, k)},$$
$$\liminf_{t \to +\infty} \frac{F_t}{\sqrt{\ln_2(t)}} = -\sqrt{2\widetilde{\sigma}(x_0, k)},$$

où $\ln_2(t) = \ln(\ln(t))$, pour tout $t > e$.
- *Si $\widetilde{W}(x_0) < k + 2\underline{W}(x_0)$, on a*

$$\lim_{t \to +\infty} F_t = \underline{m}(x_0, k).$$

On ne prouvera pas ce dernier résultat car son approche est similaire à celle employée pour montrer le Théorème 2 et il suffit ensuite de reprendre la preuve de Stout de la loi du logarithme itéré pour les martingales (cf. Stout, 1970, et aussi Meyer, 1972). De plus, il n'est peut-être pas très pertinent en théorie du recuit simulé, mais surtout, comme me l'ont fait remarquer le referee et un éditeur, il ne met pas en évidence de nouveaux phénomènes outre ceux qui apparaissent déjà dans le Théorème 2.

En fait, ces théorèmes sont aussi valables si $W(x_0) = 0$ (auquel cas on a nécessairement $\widetilde{M} = \{x_0\}$ et $\widetilde{W}(x_0) < k + 2\underline{W}(x_0)$, mis à part dans le cas trivial où $M = \{x_0\}$); il faut alors considérer les fluctuations renormalisées $F_t = \ln^{-1}(1+t)(G_t(\phi) - g_t(\phi))$ et il existe un nombre $\underline{m}(x_0, k) > 0$ tel que

$$\lim_{t \to +\infty} F_t = \underline{m}(x_0, k),$$

Si $W(x_0) < 0$, les fluctuations $F_t = G_t(\phi) - g_t(\phi)$ convergent p.s.

Par ailleurs, remarquons que l'on a pour tout $x_0 \in \widetilde{M}$ fixé,

$$\max_{x \notin \widetilde{M}} [e^{\widetilde{M}}(p_{x_0,x}) - 2U(x)] < k - 2 \min_{x \notin \widetilde{M}} U(x)$$

ainsi dans les considérations précédentes, on peut remplacer $\widetilde{W}(x_0)$ par

$$\widetilde{W}'(x_0) = \max_{x \in \widetilde{M} \setminus \{x_0\}} [e^{\widetilde{M}}(p_{x_0,x}) - 2U(x)] - 2U(x_0).$$

et il est facile de vérifier sur des exemples que le fait que $\widetilde{W}(x_0) > k + \underline{W}(x_0)$ (ou de manière équivalente que $\widetilde{W}'(x_0) > k + \underline{W}(x_0)$) et donc le type de comportement asymptotique des fluctuations, peut effectivement dépendre du point $x_0 \in \widetilde{M}$.

D'autre part, notons que des théorèmes asymptotiques pour les chaînes de Markov inhomogènes ont été obtenus par de nombreux auteurs (Iosifescu et Theodorescu, 1969 pour le théorème ergodique; Dobrushin (1956) pour le théorème de la limite centrale et Sarymsakov (1961) pour la loi du logarithme itéré), sous des hypothèses très générales, faisant intervenir le coefficient ergodique (fort de Dobrushin ou faible de Sarymsakov) des matrices de transitions, mais qui ne sont pas suffisamment précises pour la situation spécifique qui nous intéresse. Ainsi par exemple, Gantert a appliqué les résultats de (Iosifescu et Theodorescu, 1969) au recuit simulé (rectifiant par là un résultat incorrect de Gidas (1985), remarquons que le contre-exemple donné suggère qu'il n'est pas possible que $g_t(\phi)^{-1}G_t(\phi)$ converge en probabilité si ϕ est l'indicatrice d'un point $x_0 \in \widehat{M} \setminus \widetilde{M}$) pour obtenir une loi faible (respectivement forte) des grand nombres, pour $k > \gamma_0^{-1}$ (respectivement $k > 2\gamma_0^{-1}$), mais avec une constante γ_0^{-1} beaucoup plus grande que $c(M,U)$.

L'idée de la démonstration de ces théoremes est très simple :

Pour tout $s \geqslant 0$, il existe une unique fonction ψ_s définie sur M telle que

$$\begin{aligned} \phi(\cdot) - \mu_{\beta_s}(\phi) &= L_{\beta_s}\psi_s(\cdot), \\ \mu_{\beta_s}(\psi_s) &= 0. \end{aligned} \tag{1}$$

Cette fonction ψ_s est parfois appelée le potentiel associé à ϕ (par rapport à L_{β_s}, dont on aura remarqué que μ_{β_s} est l'unique probabilité invariante), et elle joue un rôle fondamental dans la théorie des chaînes de Markov finies (voir par exemple Kemeny et Snell, 1960; Iosifescu, 1980). Son intérêt pour nous provient de la formule suivante:

$$\begin{aligned} G_t(\phi) &= \int_0^t \mu_{\beta_s}(\phi)\,\mathrm{d}s + \int_0^t L_{\beta_s}\psi_s(X_s)\,\mathrm{d}s \\ &= \int_0^t \mu_{\beta_s}(\phi)\,\mathrm{d}s + \psi_t(X_t) - \psi_0(X_0) - \int_0^t \partial_s\psi_s(X_s)\,\mathrm{d}s - M_t^{\psi} \end{aligned} \tag{2}$$

où $(M_t^{\psi})_{t\geqslant 0}$ est une martingale dont le crochet oblique $\langle M^{\psi}\rangle$ est donné (p.s.) par

$$\forall\ t \geqslant 0, \qquad \langle M^{\psi}\rangle_t = \int_0^t \Gamma_{\beta_s}(\psi_s,\psi_s)(X_s)\,\mathrm{d}s$$

$\Gamma_{\beta_s}(\cdot,\cdot)$ étant le carré du champ associé à L_{β_s}; pour toute fonction ψ définie sur M et tout $x \in M$,

$$\begin{aligned} \Gamma_{\beta_s}(\psi,\psi)(x) &= L_{\beta_s}\psi^2(x) - 2\psi(x)L_{\beta_s}\psi(x) \\ &= \sum_{y\in M}(\psi(y) - \psi(x))^2 q_{\beta_s}(x,y). \end{aligned}$$

C'est à partir de l'expression (2) que l'on obtiendra le Théorème 1. En effet, une fois estimée la norme sup de ψ_t en fonction de β_t, il ne reste plus (ou presque) qu'à utiliser des résultats connus sur les martingales.

Mais pour les Théorèmes 2 et 3, l'estimée de la norme sup de ψ_t ne suffit plus, c'est pourquoi on a dû se restreindre aux arbres, situation dans laquelle on sait expliciter sous une forme particulièrement simple les potentiels ψ_s. L'hypothèse $\widehat{M} = \widetilde{M}$ permettra

d'appliquer le Théorème ergodique (ou plutôt l'une de ses variantes, voir la Proposition 9 ci-dessous) en tout point de $\widehat{M}$. On prouvera notamment qu'il existe des constantes
$\widehat{W}(x_0,k)$ et $\sigma(x_0,k) > 0$ telles qu'en temps grand, p.s.

$$\begin{cases} \langle M^{\psi}\rangle_t \text{ converge} & \text{si } \widehat{W}(x_0,k) < -k, \\ \langle M^{\psi}\rangle_t \sim \sigma(x_0,k)\ln(1+t) & \text{si } \widehat{W}(x_0,k) = -k, \\ \langle M^{\psi}\rangle_t \sim \sigma(x_0,k)t^{1+k^{-1}\widehat{W}(x_0,k)} & \text{si } \widehat{W}(x_0,k) > -k, \end{cases}$$

mais il se pourra que $\widehat{W}(x_0,k) \neq \frac{1}{2}(1+k^{-1}\widetilde{W}(x_0))$ et $\widehat{W}(x_0,k) \neq 1+k^{-1}\underline{W}(x_0)$.

Le plan de l'article est le suivant: Dans la section suivante on estimera les potentiels ψ_s et on en déduira le Théorème ergodique pour les indicatrices d'éléments de $\widetilde{M}$. Le cas des éléments $x_0 \in \widehat{M}$ tels que $U(x_0) = k$ sera traité dans la Section 3 et on y montrera également que tous les points de $\widehat{M}$ sont récurrents et que tous les autres sont transcients. Ensuite, dans la Section 4, on explicitera les potentiels ψ_s dans le cas d'un arbre, ce qui nous permettra de démontrer le Théorème 2 dans la cinquième section. Enfin, on verra dans la dernière section que le Théorème 1 s'étend immédiatement aux cas où on fait seulement l'hypothèse de réversibilité faible de Hajek.

Par ailleurs, signalons que comportement des temps d'occupations convenablement renormalisés en les points de $\widehat{M}\setminus\widehat{M}$ a été résolu dans (Miclo 1994) que nous venons de soumettre pour publication: On obtient la convergence étroite vers la loi à l'instant 1 d'un certain processus de sauts à accroissements indépendants.

2. Ergodicité des points de $\widetilde{M}$

Rappelons l'interprétation probabiliste du potentiel ψ_s solution de (1) (où ϕ est l'indicatrice d'un point $x_0 \in M$ fixé, et où $s \geqslant 0$ est également fixé). Soit $(Z_t^{x,s})_{t\geqslant 0}$ un processus de Markov (càdlàg.) homogène à valeurs dans M, issu de $x \in M$ et dont le générateur est L_{β_s}. Alors,

$$\psi_s(x) = \boldsymbol{E}\left[\int_0^\infty \phi(Z_t^{x,s}) - \mu_{\beta_s}(\phi)\,\mathrm{d}t\right].$$

Notons $m_t^{x,s}$ la loi de $Z_t^{x,s}$, pour $t \geqslant 0$. Il est bien connu (ceci découle par exemple de l'étude de l'évolution de l'entropie de $m_t^{x,s}$ par rapport à μ_{β_s}, à l'aide d'inégalités de Sobolev logarithmiques et en utilisant le fait qu'au temps $t = 0$ cette entropie est majorée par $K_1(\beta_s \vee 1)$ pour une certaine constante $K_1 > 0$ indépendante de $s \geqslant 0$, voir (Miclo, 1992) qu'il existe deux constantes $K_2 > 0$ et $p > 0$, telles que pour tout $(x,s) \in M \times [0,+\infty[$ et tout $t \geqslant 0$,

$$\|m_t^{x,s} - \mu_{\beta_s}\|_{\mathrm{VT}} \leqslant K_2(\beta_s \vee 1)^p \exp(-K_2^{-1}\exp[-\beta_s c(M,U)]t)$$

où $\| \ \|_{\mathrm{VT}}$ désigne la variation totale. Ainsi, pour tout $x \in M$,

$$\begin{aligned}|\psi_s(x)| &= \left|\boldsymbol{E}\left[\int_0^\infty \phi(Z_t^{x,s}) - \mu_{\beta_s}(\phi)\,\mathrm{d}t\right]\right| \\ &= \left|\int_0^\infty (m_t^{x,s} - \mu_{\beta_s})(\phi)\,\mathrm{d}t\right| \\ &\leqslant \int_0^\infty \|m_t^{x,s} - \mu_{\beta_s}\|_{\mathrm{VT}}\,\mathrm{d}t \\ &\leqslant K_2^2(\beta_s \vee 1)^p \exp[\beta_s c(M,U)].\end{aligned}$$

Pour $\psi \in F(M)$ (qui est l'ensemble des fonctions réelles sur M), on note

$$\|\psi\| = \max_{z\in M} |\psi(z)|\,.$$

On vient donc de voir que pour tout $s \geqslant 0$,

$$\|\psi_s\| \leqslant K_2^2(\beta_s \vee 1)^p \exp[\beta_s c(M,U)].$$

Cependant, mis à part pour le cas où $x_0 \in N$, le facteur $c(M,U)$ après β_s dans l'exponentielle n'est pas optimal (voir les calculs explicites effectués dans la Section 4, si (M,q) est un arbre). Notre but dans un premier temps est d'obtenir la majoration suivante, qui ensuite sera légèrement améliorée (voir la formule (4) ci-dessous):

Proposition 4. *Soit* $V(x_0) = \max_{x\in M}[H(x_0,x) - U(x) - U(x_0)] \geqslant 0$. *Il existe deux constantes* $K_3 > 0$ *et* $\bar{p} > 0$, *telles que pour tout* $s \geqslant 0$,

$$\|\psi_s\| \leqslant K_3(\beta_s \vee 1)^{\bar{p}} \exp[\beta_s V(x_0)]. \tag{4}$$

Démonstration. Il est facile de vérifier que pour tout $z_0 \in N$ fixé, $c(M,U) = \max_{x\in M} H(z_0,x) - U(x) - U(z_0)$. Ainsi, si $x_0 \in N$, $V(x_0) = c(M,U)$ et la proposition est établie.

Si $x_0 \notin N$, il faut rajouter un point au graphe : Soit $\bar{x}_0 \notin M$ et $\overline{M} = M \sqcup \{\bar{x}_0\}$. On considère le noyau de mesures (d'intensités) de transitions $\bar{q}$ défini sur $\overline{M} \times \overline{M}$ par

$$\bar{q}(x,y) = \begin{cases} q(x,y) & \text{si } x,y \in M, \\ 1 & \text{si } x = \bar{x}_0 \text{ et } y = x_0, \\ 1 & \text{si } x = x_0 \text{ et } y = \bar{x}_0, \\ 0 & \text{sinon.} \end{cases}$$

et la probabilité $\bar{\mu}$ définie sur $\overline{M}$ par

$$\bar{\mu}(x) = \begin{cases} (1 + \mu(x_0))^{-1}\mu(x) & \text{si } x \in M, \\ (1 + \mu(x_0))^{-1}\mu(x_0) & \text{si } x = \bar{x}_0. \end{cases}$$

Il est clair que $\bar{q}$ est réversible par rapport à $\bar{\mu}$. Soit également le potentiel $\bar{U}$ défini sur $\overline{M}$ par

$$\bar{U}(x) = \begin{cases} U(x) & \text{si } x \in M, \\ 0 & \text{si } x = \bar{x}_0. \end{cases}$$

Comme dans l'introduction, on considère pour tout $s \geqslant 0$, le noyau de mesures de transitions $\bar{q}_{\beta_s}$ donné par

$$\forall\ x, y \in \overline{M}, \qquad \bar{q}_{\beta_s}(x, y) = \begin{cases} \exp(-\beta_s(\bar{U}(y) - \bar{U}(x))_+)\bar{q}(x, y) & \text{si } y \neq x, \\ 1 - \sum_{z \in \overline{M} \setminus \{x\}} \bar{q}_{\beta_s}(x, z) & \text{si } y = x. \end{cases}$$

l'opérateur $\bar{L}_{\beta_s}$ associé sur $F(\overline{M})$,

$$\forall\ \phi \in F(\overline{M}),\ \forall\ x \in \overline{M}, \qquad \bar{L}_{\beta_s}\phi(x) = \sum_{y \in \overline{M}} (\phi(y) - \phi(x))\bar{q}_{\beta_s}(x, y)$$

et la probabilité de Gibbs $\bar{\mu}_{\beta_s}$ invariante pour cet opérateur,

$$\forall\ x \in \overline{M}, \qquad \bar{\mu}_{\beta_s}(x) = \bar{Z}_{\beta_s}^{-1} \exp(-\beta_s \bar{U}(x))\bar{\mu}(x)$$

où $\bar{Z}_{\beta_s}$ est la constante de normalisation.

On s'intéresse à l'unique solution $\bar{\psi}_s \in F(\overline{M})$ de l'équation

$$\begin{aligned} \bar{\phi}(\cdot) - \bar{\mu}_{\beta_s}(\bar{\phi}) &= \bar{L}_{\beta_s}\bar{\psi}_s(\cdot), \\ \bar{\mu}_{\beta_s}(\bar{\psi}_s) &= 0 \end{aligned}$$

où $\bar{\phi}$ est l'indicatrice de $\bar{x}_0$.

Soit $(\bar{Z}_t^{x,s})_{t \geqslant 0}$ un processus de Markov (càdlàg.) homogène à valeurs dans $\overline{M}$, issu de $x \in \overline{M}$ et dont le générateur est $\bar{L}_{\beta_s}$. Alors,

$$\bar{\psi}_s(x) = \boldsymbol{E}\left[\int_0^\infty \bar{\phi}(\bar{Z}_t^{x,s}) - \bar{\mu}_{\beta_s}(\bar{\phi})\,\mathrm{d}t\right]$$

et les considérations précédentes permettent de voir qu'il existe deux constantes $K_4 > 0$ et $\bar{p} > 0$, telles que pour tout $s \geqslant 0$,

$$\|\bar{\psi}_s\|_{\overline{M}} \leqslant K_4(\beta_s \vee 1)^{\bar{p}} \exp[\beta_s c(\overline{M}, \bar{U})],$$

$\|\ \|_{\overline{M}}$ désignant la norme sup sur $\overline{M}$. Cependant, en utilisant le fait que $\bar{x}_0$ est un minimum global de $\bar{U}$ que l'on ne peut atteindre qu'en passant par x_0, il est clair que

$$c(\overline{M}, \bar{U}) = \max_{x \in M} \min_{p \in \mathscr{C}_{x_0,x}} e(p) - U(x)$$

ainsi,

$$\|\bar{\psi}_s\|_{\overline{M}} \leqslant K_4(\beta_s \vee 1)^{\bar{p}} \exp(\beta_s[V(x_0) + U(x_0)]).$$

Mais on sait exprimer ψ_s en fonction de $\bar{\psi}_s$. En effet, calculons $L_{\beta_s}\bar{\psi}_{s,M}$, où $\bar{\psi}_{s,M}$ est la restriction de $\bar{\psi}_s$ à M:

– Pour $x \in M$, $x \neq x_0$,

$$\begin{aligned} L_{\beta_s}\bar{\psi}_{s,M}(x) &= \bar{L}_{\beta_s}\bar{\psi}_{s,M}(x) \\ &= -\bar{\mu}_{\beta_s}(\bar{\phi}) = -\bar{\mu}_{\beta_s}(\bar{x}_0) \end{aligned}$$

remarquons d'ailleurs que

$$\begin{aligned}\bar{\mu}_{\beta_s}(\bar{x}_0) &= \bar{Z}_{\beta_s}^{-1}\bar{\mu}(\bar{x}_0)\\ &\leqslant \frac{\mu(x_0)}{(1+\mu(x_0))\bar{\mu}(N \sqcup \{\bar{x}_0\})}\\ &= \frac{\mu(x_0)}{\mu(N)+\mu(x_0)},\end{aligned}$$

ce qui montre que

$$(1-\bar{\mu}_{\beta_s}(\bar{x}_0))^{-1} \leqslant \frac{\mu(N)+\mu(x_0)}{\mu(N)}.$$

– Si $x = x_0$,

$$\begin{aligned}L_{\beta_s}\bar{\psi}_{s,M}(x_0) &= \bar{L}_{\beta_s}\bar{\psi}_{s,M}(x_0) - (\bar{\psi}_s(\bar{x}_0) - \bar{\psi}_s(x_0))\bar{q}_{\beta_s}(x_0,\bar{x}_0)\\ &= -\bar{\mu}_{\beta_s}(\bar{\phi}) + \bar{L}_{\beta_s}\bar{\psi}_{s,M}(\bar{x}_0)\bar{q}_{\beta_s}(\bar{x}_0,x_0)^{-1}\bar{q}_{\beta_s}(x_0,\bar{x}_0)\\ &= (1-\bar{\mu}_{\beta_s}(\bar{x}_0))\exp(\beta_s U(x_0)) - \bar{\mu}_{\beta_s}(\bar{x}_0).\end{aligned}$$

Il apparaît ainsi que

$$L_{\beta_s}\bar{\psi}_{s,M} = (1-\bar{\mu}_{\beta_s}(\bar{x}_0))\exp(\beta_s U(x_0))\phi - (1-\bar{\mu}_{\beta_s}(\bar{x}_0))\exp(\beta_s U(x_0))\mu_{\beta_s}(\phi).$$

c'est-à-dire que $(1-\bar{\mu}_{\beta_s}(\bar{x}_0))\exp(\beta_s U(x_0))\psi_s$ et $\bar{\psi}_{s,M}$ ne diffèrent que par une constante additive. Celle-ci vaut d'ailleurs

$$\begin{aligned}\mu_{\beta_s}(\bar{\psi}_{s,M}) &= \frac{\bar{Z}_{\beta_s}}{Z_{\beta_s}}(1+\mu(x_0))\bar{\mu}_{\beta_s}(\bar{\psi}_s 1_M)\\ &= \frac{\bar{Z}_{\beta_s}}{Z_{\beta_s}}(1+\mu(x_0))[\bar{\mu}_{\beta_s}(\bar{\psi}_s) - \bar{\mu}_{\beta_s}(\bar{x}_0)\bar{\psi}_s(\bar{x}_0)]\\ &= -Z_{\beta_s}^{-1}\mu(x_0)\bar{\psi}_s(\bar{x}_0),\end{aligned}$$

d'où

$$\psi_s = (1-\bar{\mu}_{\beta_s}(\bar{x}_0))^{-1}\exp(-\beta_s U(x_0))[\bar{\psi}_{s,M} - Z_{\beta_s}^{-1}\mu(x_0)\bar{\psi}_s(\bar{x}_0)],$$

puis

$$\begin{aligned}\|\psi_s\| &\leqslant (1-\bar{\mu}_{\beta_s}(\bar{x}_0))^{-1}\exp(-\beta_s U(x_0))(1+Z_{\beta_s}^{-1}\mu(x_0))\,\|\bar{\psi}_s\|_{\overline{M}}\\ &\leqslant \frac{\mu(N)+\mu(x_0)}{\mu(N)}(1+\mu(N)^{-1}\mu(x_0))\exp(-\beta_s U(x_0))\,\|\bar{\psi}_s\|_{\overline{M}}\\ &= (1+\mu(N)^{-1}\mu(x_0))^2\exp(-\beta_s U(x_0))\,\|\bar{\psi}_s\|_{\overline{M}}\end{aligned}$$

et le résultat annoncé.

Notons que $V(x_0)$ est positif; en effet,

$$V(x_0) = \max_{x\in M}[\min_{p\in\mathscr{C}_{x_0,x}} e(p) - U(x) - U(x_0)]$$

ainsi c'est clair si $x_0 \in N$, et sinon il suffit de prendre $x \in N$ ci-dessus pour le voir.

Remarquons également qu'en fait $V(x_0) = V'(x_0) \stackrel{\text{déf}}{=} \max_{x \neq x_0}[H(x_0,x) - U(x) - U(x_0)]$, mis à part dans la situation triviale où $M = \{x_0\}$, auquel cas la majoration (3) est aussi satisfaite avec $V'(x_0)$, puisqu'alors $\psi_s \equiv 0$ et $V'(x_0) = -\infty$. $\square$

Par ailleurs, il est bien connu (voir par exemple le début de la Section 4) que $\psi_s(x)$ s'exprime comme une fraction rationnelle F_x (dépendant aussi du x_0 fixé) en les variables $(q_s(y,z))_{y,z \in M}$. Ceci permet de voir qu'il existe des entiers $h,k \geqslant 0$ et des réels $a_0, a_1, \cdots, a_h$, $1 = a'_0, a'_1, \ldots, a'_k$, $0 \geqslant \delta_0 > \delta_1 > \ldots > \delta_h$ et $0 = \delta'_0 > \delta'_1 > \delta'_2 > \cdots > \delta'_h$ (tous ces nombres dépendant de x et de x_0), tels que pour tout $s \geqslant 0$,

$$\psi_s(x) = \exp(\beta_s V(x_0)) \frac{\sum_{i=1}^h a_i \exp(\beta_s \delta_i)}{\sum_{j=1}^k a'_j \exp(\beta_s \delta'_j)}.$$

Ceci permet de voir qu'en fait, il existe une constante $K_5 > 0$ telle que pour tout $s \geqslant 0$,

$$\|\psi_s\| \leqslant K_5 \exp[\beta_s V(x_0).] \tag{5}$$

De plus, vu la forme de β_s, on a

$$\partial_s \psi_s(x) = k^{-1} \left[V(x_0) + \frac{\sum_{i=1}^h a_i \delta_i \exp(\beta_s \delta_i)}{\sum_{i=1}^h a_i \exp(\beta_s \delta_i)} - \frac{\sum_{j=1}^k a'_j \delta'_j \exp(\beta_s \delta'_j)}{\sum_{j=1}^k a'_j \exp(\beta_s \delta'_j)} \right] \frac{1}{1+s} \psi_s(x)$$

ce qui permet d'obtenir :

Proposition 5. *Il existe des constantes $K_6 > 0$ et $\delta > 0$ telles que pour tout $s \geqslant 0$,*
– si $V(x_0) > 0$, $\|\partial_s \psi_s\| \leqslant K_6 \exp[\beta_s(V(x_0) - k)]$,
– si $V(x_0) = 0$, $\|\partial_s \psi_s\| \leqslant K_6 \exp[-\beta_s(k + \delta)]$.

Nous disposons maintenant de toutes les estimées nécessaires à la démonstration du Théorème ergodique pour les points de $\widetilde{M}$. Notons que jusqu'à présent, on n'a pas eu besoin de l'hypothèse $x_0 \in \widetilde{M}$, les Propositions 4 et 5 étant valables pour tout $x_0 \in M$.

Supposons désormais que $x_0 \in \widetilde{M}$. On va comparer chacun des termes de

$$G_t(\phi) = \int_0^t \mu_{\beta_s}(\phi)\, ds + \psi_t(X_t) - \psi_0(X_0) - \int_0^t \partial_s \psi_s(X_s)\, ds - M_t^{\psi}$$

pour obtenir que le premier est dominant.

Du fait que $U(x_0) < k$, on a vu dans l'introduction qu'en temps grand,

$$\int_0^t \mu_{\beta_s}(\phi)\, ds \sim \mu(N)^{-1} \mu(x_0)(1 - k^{-1} U(x_0))^{-1} \exp(\beta_t[k - U(x_0)])$$

ainsi, on a déjà

$$\int_0^t \mu_{\beta_s}(\phi)\, ds \gg \psi_0(X_0).$$

Montrons que

$$\int_0^t \mu_{\beta_s}(\phi)\,\mathrm{d}s \gg \|\psi_t\| \geqslant |\psi_t(X_t)|. \tag{6}$$

Il suffit pour cela, d'après la formule (4), de vérifier que $V(x_0) < k - U(x_0)$, ou encore que pour tout $x \in M$,

$$\exists\ p \in \mathscr{C}_{x,x_0},\ \text{tel que } e(p) - U(x) < k. \tag{7}$$

Distinguons deux cas :

- si $x \in \widetilde{M}$, alors, par définition de $\widetilde{M}$ et puisque $x_0 \in \widetilde{M}$, il existe un chemin $p \in \mathscr{C}_{x,x_0}$ tel que $e(p) < k$, d'où (6).
- si $x \notin \widetilde{M}$, alors $H(x,x_0) \geqslant k$, du fait que $x_0 \in \widetilde{M}$. Cependant, soit $z_0 \in N$, puisque $H(x_0,z_0) < k$, il est clair que $H(x,x_0) = H(x,z_0)$. Ainsi, il existe un chemin $p \in \mathscr{C}_{x,x_0}$ tel que

$$\begin{aligned} e(p) - U(x) &= H(x,z_0) - U(x) - U(z_0) \\ &\leqslant c(M,U) \\ &< k \end{aligned}$$

d'où (6) puis finalement (5).

De la même manière, en utilisant la Proposition 5, on montre que

$$\int_0^t \mu_{\beta_s}(\phi)\,\mathrm{d}s \gg \int_0^t \|\partial_s\psi_s\|\,\mathrm{d}s \geqslant \left|\int_0^t \partial_s\psi_s(X_s)\,\mathrm{d}s\right|.$$

Il reste donc à s'intéresser au comportement de la partie martingale M_t^ψ. On va utiliser, pour l'étudier, le résultat suivant présenté par Lepingle dans (Lepingle, 1978).

Proposition 6. *Soit $(M_t)_{t\geqslant 0}$ une martingale localement de carré intégrable (par rapport à un espace probabilisé filtré $(\Omega,\mathscr{F},(\mathscr{F}_t)_{t\geqslant 0},\boldsymbol{P})$ satisfaisant aux conditions habituelles). On peut donc définir son crochet oblique $(\langle M\rangle_t)_{t\geqslant 0}$. Alors, p.s.,*

– Sur $\{\langle M\rangle_\infty < +\infty\}$, M_t converge en temps grand,

– Sur $\{\langle M\rangle_\infty = +\infty\}$, on a en temps grand,

$$|M_t| \ll \ln(\langle M\rangle_t)\sqrt{\langle M\rangle_t}.$$

Appliquons cette proposition avec la martingale localement bornée $(M_t)_{t\geqslant 0} = (M_t^\psi)_{t\geqslant 0}$:

- Si $\langle M^\psi\rangle_\infty < +\infty$, puisque M_t^ψ converge, on a bien

$$\int_0^t \mu_{\beta_s}(\phi)\,\mathrm{d}s \gg \left|M_t^\psi\right|$$

(mais on vérifiera a posteriori dans la Section 5 que p.s., en général, on n'a pas $\langle M^\psi\rangle_\infty < +\infty$).

• Le cas intéressant est donc si $\langle M^{\psi}\rangle_{\infty} = +\infty$. On a vu que pour tout $t \geqslant 0$,

$$\begin{aligned}\langle M^{\psi}\rangle_t &= \int_0^t \Gamma_{\beta_s}(\psi_s, \psi_s)(X_s)\,\mathrm{d}s\\ &= I_1(t) + I_2(t)\end{aligned}$$

où on a posé :

$$\begin{aligned}I_1(t) &= \int_0^t L_{\beta_s}\psi_s^2(X_s)\,\mathrm{d}s\\ I_2(t) &= -2\int_0^t \psi_s(X_s)L_{\beta_s}\psi_s(X_s)\,\mathrm{d}s\end{aligned}$$

Considérons d'abord la première intégrale

$$I_1(t) = \psi_t^2(X_t) - \psi_0^2(X_0) - \int_0^t \partial_s\psi_s^2(X_s)\,\mathrm{d}s - M_t^{\psi^2}$$

où $(M_t^{\psi^2})_{t\geqslant 0}$ est une martingale dont le crochet oblique est donné par

$$\forall\ t \geqslant 0, \qquad \left\langle M^{\psi^2}\right\rangle_t = \int_0^t \Gamma_{\beta_s}(\psi_s^2, \psi_s^2)(X_s)\,\mathrm{d}s$$

On majore les trois premiers termes dans la décomposition précédente de $I_1(t)$ en utilisant l'estimée (4) et la Proposition 5:

$$\begin{aligned}\left|\psi_t^2(X_t) - \psi_0^2(X_0) - \int_0^t \partial_s\psi_s^2(X_s)\,\mathrm{d}s\right| &\leqslant \|\psi_t\|^2 + \|\psi_0\|^2 + 2\int_0^t \|\psi_s\|\,\|\partial_s\psi_s\|\,\mathrm{d}s\\ &\leqslant K_7\exp(2\beta_t V(x_0))\end{aligned}$$

pour une certaine constante $K_7 \geqslant 0$, indépendante de t.

– Si $\left\langle M^{\psi^2}\right\rangle_{\infty} < +\infty$, d'après la Proposition 6, $M_t^{\psi^2}$ converge en temps grand, et il existe donc une constante (aléatoire) $K_8 \geqslant K_7$, telle que pour tout $t \geqslant 0$,

$$|I_1(t)| \leqslant K_8\exp(2\beta_t V(x_0)).$$

– Si $\left\langle M^{\psi^2}\right\rangle_{\infty} = +\infty$, toujours d'après la Proposition 6, il existe une nouvelle constante $K_9 \geqslant 0$ (désormais, les K_i, $i \geqslant 9$, désigneront des constantes indépendantes du temps mais éventuellement aléatoires, comme cette dernière), telle que pour tout $t \geqslant 0$,

$$\left|M_t^{\psi^2}\right| \leqslant K_9\ln\left(e \vee \left\langle M^{\psi^2}\right\rangle_t\right)\sqrt{1 + \langle M^{\psi^2}\rangle_t}.$$

Pour majorer $\ln(e \vee \left\langle M^{\psi^2} \right\rangle_t)$, utilisons l'estimation grossière

$$\begin{aligned}\left\langle M^{\psi^2} \right\rangle_t &\leqslant \int_0^t \|\Gamma_{\beta_s}(\psi_s^2, \psi_s^2)\| \, \mathrm{d}s \\ &\leqslant 4 \int_0^t \|\psi_s\|^4 \, \mathrm{d}s \\ &\leqslant K_{10} \exp(\beta_t[4V(x_0) + k]).\end{aligned}$$

Par ailleurs, remarquons que pour tout $t \geqslant s \geqslant 0$ et tout $x \in M$,

$$\begin{aligned}\Gamma_{\beta_s}(\psi_s^2, \psi_s^2)(x) &= \sum_{y \in M} (\psi_s^2(y) - \psi_s^2(x))^2 q_{\beta_s}(x, y) \\ &= \sum_{y \in M} (\psi_s(y) + \psi_s(x))^2 (\psi_s(y) - \psi_s(x))^2 q_{\beta_s}(x, y) \\ &\leqslant 4 \|\psi_s\|^2 \, \Gamma_{\beta_s}(\psi_s, \psi_s)(x) \\ &\leqslant K_{11} \exp(2\beta_t V(x_0)) \Gamma_{\beta_s}(\psi_s, \psi_s)(x)\end{aligned}$$

et qu'il s'en suit que

$$\left\langle M^{\psi^2} \right\rangle_t \leqslant K_{11} \exp(2\beta_t V(x_0)) \left\langle M^{\psi} \right\rangle_t.$$

Il existe donc une constante $K_{12} \geqslant 0$, telle que pour tout $t \geqslant 0$,

$$\left| M_t^{\psi^2} \right| \leqslant K_{12}(1 + \beta_t) \exp(\beta_t V(x_0)) \sqrt{1 + \langle M^{\psi} \rangle_t}$$

et ainsi, dans tous les cas, on peut écrire que pour tout $t \geqslant 0$,

$$|I_1(t)| \leqslant K_{13} \exp(2\beta_t V(x_0)) + K_{14} \beta_t \exp(\beta_t V(x_0)) \sqrt{\langle M^{\psi} \rangle_t}.$$

Intéressons-nous maintenant à la seconde intégrale $I_2(t)$, on a

$$\begin{aligned}|I_2(t)| &= 2 \left| \int_0^t \psi_s(X_s) L_{\beta_s} \psi_s(X_s) \, \mathrm{d}s \right| \\ &= 2 \left| \int_0^t \psi_s(X_s) [\phi(X_s) - \mu_{\beta_s}(\phi)] \, \mathrm{d}s \right| \\ &\leqslant 2 \int_0^t \|\psi_s\| \, \mu_{\beta_s}(\phi) \, \mathrm{d}s + 2 \int_0^t \|\psi_s\| \, \phi(X_s) \, \mathrm{d}s \\ &\leqslant K_{15} \exp(\beta_t[V(x_0) - U(x_0) + k]) + K_{16} \exp(\beta_t V(x_0)) G_t(\phi).\end{aligned}$$

En regroupant ces majorations et en utilisant l'inégalité $k - U(x_0) > V(x_0)$ que l'on a vérifiée précédemment, on obtient finalement que $\langle M^{\psi} \rangle_t$ satisfait

$$\begin{aligned}\langle M^{\psi} \rangle_t \leqslant\; & K_{14} \beta_t \exp(\beta_t V(x_0)) \sqrt{\langle M^{\psi} \rangle_t} + K_{17} \exp(\beta_t[V(x_0) \\ & - U(x_0) + k]) + K_{16} \exp(\beta_t V(x_0)) G_t(\phi)\end{aligned}$$

ce qui implique que

$$\langle M^{\psi}\rangle_t \leqslant K_{14}^2\beta_t^2 \exp(2\beta_t V(x_0)) + 2K_{17}\exp(\beta_t[V(x_0) - U(x_0) + k]) + 2K_{16}\exp(\beta_t V(x_0))G_t(\phi),$$

ou encore que

$$\langle M^{\psi}\rangle_t \leqslant K_{18}\exp(\beta_t[V(x_0) - U(x_0) + k]) + 2K_{16}\exp(\beta_t V(x_0))G_t(\phi).$$

On en déduit qu'en temps grand (rappelons que l'on est sous l'hypothèse $\langle M^{\psi}\rangle_\infty = +\infty$),

$$\left|M_t^{\psi}\right| \ll \beta_t \exp\left(\beta_t \frac{V(x_0) - U(x_0) + k}{2}\right) + \beta_t \exp\left(\beta_t \frac{V(x_0)}{2}\right)\sqrt{G_t(\phi)}.$$

Cependant, d'après ce que l'on a déjà vu,

$$\begin{aligned}
&\beta_t \exp\left(\beta_t \frac{V(x_0)}{2}\right)\sqrt{G_t(\phi)} \\
&\quad\leqslant \beta_t \exp\left(\beta_t \frac{V(x_0)}{2}\right)\sqrt{K_{19}\exp(\beta_t[k - U(x_0)]) + \left|M_t^{\psi}\right|} \\
&\quad\leqslant \beta_t \exp\left(\beta_t \frac{V(x_0)}{2}\right)\left(\sqrt{K_{19}\exp(\beta_t[k - U(x_0)])} + \sqrt{\left|M_t^{\psi}\right|}\right) \\
&\quad\leqslant K_{19}^{\frac{1}{2}}\beta_t \exp\left(\beta_t \frac{V(x_0) - U(x_0) + k}{2}\right) + \frac{1}{2}\beta_t^2 \exp(\beta_t V(x_0)) + \frac{1}{2}\left|M_t^{\psi}\right|
\end{aligned}$$

d'où en fait,

$$\left|M_t^{\psi}\right| \ll \beta_t \exp\left(\beta_t \frac{V(x_0) - U(x_0) + k}{2}\right).$$

Ainsi on obtient dans ce cas aussi

$$\int_0^t \mu_{\beta_s}(\phi)\,\mathrm{d}s \gg \left|M_t^{\psi}\right|$$

en utilisant une fois de plus l'inégalité $k - U(x_0) > V(x_0)$.

En définitive, on a donc montré que

$$\begin{aligned}
G_t(\phi) &\sim \int_0^t \mu_{\beta_s}(\phi)\,\mathrm{d}s \\
&\sim \mu(N)^{-1}\mu(x_0)(1 - k^{-1}U(x_0))^{-1}t^{(1-k^{-1}U(x_0))}
\end{aligned}$$

d'où le Théorème 1 dans le cas où $x_0 \in \widetilde{M}$.

3. D'autres résultats d'ergodicité

Une démonstration identique à celle de la section précédente permet d'obtenir le Théorème ergodique en les points $x_0 \in \widehat{M}\backslash\widetilde{M}$. Mais si on veut améliorer la connaissance que l'on a de la vitesse de convergence, on peut remplacer la Proposition 6 par la loi du logarithme itéré qui est plus précise ; on utilisera donc plutôt le résultat suivant, également présenté par Lepingle dans (Lepingle 1978) (qui en donne en fait une version moins restrictive, voir son Théorème 3).

Proposition 7. *Soit $(M_t)_{t\geqslant 0}$ une martingale localement de carré intégrable. On note $\triangle M_t = M_t - M_{t-}$ le saut de cette martingale au temps $t\geqslant 0$, et on suppose qu'il existe $K_1 > 0$ tel que pour tout $t\geqslant 0$, $|\triangle M_t| \leqslant K_1$. Alors, p.s. sur $\{\langle M\rangle_\infty = +\infty\}$,*

$$\limsup_{t\to+\infty} \frac{M_t}{\sqrt{2\langle M\rangle_t \ln_2(\langle M\rangle_t)}} \leqslant 1.$$

Plus précisément, on va montrer un résultat un peu plus général que la loi des grands nombres pour les points $x_0 \in \widehat{M}\setminus\widetilde{M}$: Pour ϕ l'indicatrice d'un point $x_0 \in \widehat{M}$, notons

$$\widetilde{G}_t(\phi) = \int_0^t \exp(\beta_s[U(x_0) - k])\phi(X_s)\,\mathrm{ds}$$

et

$$\widetilde{g}_t(\phi) = \int_0^t \exp(\beta_s[U(x_0) - k])\mu_{\beta_s}(\phi)\,\mathrm{ds}$$

Proposition 8. *Pour tout $x_0 \in \widehat{M}$, on a p.s.,*

$$\lim_{t\to+\infty} \widetilde{g}_t^{-1}(\phi)\widetilde{G}_t(\phi) = 1.$$

Cette proposition termine la démonstration du Théorème 1, car pour $x_0 \in \widehat{M}\setminus\widetilde{M}$, on a $U(x_0) = k$, ainsi $\widetilde{G}_t(\phi) = G_t(\phi)$ et $\widetilde{g}_t(\phi) = g_t(\phi)$.

Démonstration. Comme annoncé, il suffit de reprendre la démonstration de la section précédente en utilisant la Proposition 7. Pour $s\geqslant 0$, soit $\widetilde{\psi}_s = \exp(\beta_s[U(x_0) - k])\psi_s$, cette fonction est l'unique solution de

$$\begin{cases} \exp(\beta_s[U(x_0) - k])\phi(\cdot) - \exp(\beta_s[U(x_0) - k])\mu_{\beta_s}(\phi) = L_{\beta_s}\widetilde{\psi}_s(\cdot) \\ \mu_{\beta_s}(\widetilde{\psi}_s) = 0. \end{cases}$$

Ainsi,

$$\widetilde{G}_t(\phi) = \widetilde{g}_t(\phi) + \widetilde{\psi}_t(X_t) - \widetilde{\psi}_0(X_0) - \int_0^t \partial_s\widetilde{\psi}_s(X_s)\,\mathrm{ds} - M_t^{\widetilde{\psi}}, \tag{8}$$

où $(M_t^{\widetilde{\psi}})_{t\geqslant 0}$ est une martingale dont le crochet oblique $\left\langle M^{\widetilde{\psi}}\right\rangle$ est donné par

$$\forall\ t\geqslant 0, \qquad \left\langle M^{\widetilde{\psi}}\right\rangle_t = \int_0^t \Gamma_{\beta_s}(\widetilde{\psi}_s, \widetilde{\psi}_s)(X_s)\,\mathrm{d}s.$$

Notons que $\widetilde{g}_t(\phi)$ diverge en temps grand et que

$$\widetilde{g}_t(\phi) \sim \mu(N)^{-1}\mu(x_0)\ln(1+t).$$

D'autre part, on vérifie que $x_0 \in \widehat{M}$ implique que $V(x_0) \leqslant k - U(x_0)$, il existe donc une constante $K_2 > 0$ telle que pour tout $t \geqslant 0$,

$$\left\|\widetilde{\psi}_t\right\| \leqslant K_2$$

et de la même manière que l'on a montré la Proposition 6, on voit qu'il existe deux constantes $K_3 > 0$ et $\delta > 0$ telles que pour tout $t \geqslant 0$,

$$\left\|\partial_t \widetilde{\psi}_t\right\| \leqslant K_3 \exp(-\beta_t[k+\delta]).$$

Ceci permet de voir que l'expression

$$\widetilde{\psi}_t(X_t) - \widetilde{\psi}_0(X_0) - \int_0^t \partial_s \widetilde{\psi}_s(X_s)\,\mathrm{d}s$$

reste bornée et donc notamment qu'en temps grand,

$$\widetilde{g}_t(\phi) \gg \left|\widetilde{\psi}_t(X_t) - \widetilde{\psi}_0(X_0) - \int_0^t \partial_s \widetilde{\psi}_s(X_s)\,\mathrm{d}s\right|.$$

Ainsi, le seul comportement qui peut éventuellement poser problème est celui de $M^{\widetilde{\psi}}$. Remarquons que les sauts de cette martingale sont donnés par $\triangle M_t^{\widetilde{\psi}} = \triangle \widetilde{\psi}_t(X_t)$, et qu'ils sont donc bornés (par $2K_2$). D'après la Proposition 7, il existe alors une constante $K_4 > 0$ (aléatoire) telle que pour tout t, on ait

$$\left|M_t^{\widetilde{\psi}}\right| \leqslant K_4 + 2\sqrt{\left\langle M^{\widetilde{\psi}}\right\rangle_t \ln_2\left(\left\langle M^{\widetilde{\psi}}\right\rangle_t + \mathrm{e}^e\right)}.$$

Pour estimer $\left\langle M^{\widetilde{\psi}}\right\rangle_t$, on procède comme précédemment; pour $t \geqslant 1$,

$$\begin{aligned}
\left\langle M^{\widetilde{\psi}}\right\rangle_t &= \widetilde{\psi}_t^2(X_t) - \widetilde{\psi}_0^2(X_0) - 2\int_0^t \widetilde{\psi}_s(X_s)(\partial_s\widetilde{\psi}_s)(X_s)\,\mathrm{d}s - M_t^{\widetilde{\psi}^2} \\
&\quad -2\int_0^t \widetilde{\psi}_s(X_s)\exp(\beta_s[U(x_0)-k])\phi(X_s)\,\mathrm{d}s \\
&\quad +2\int_0^t \widetilde{\psi}_s(X_s)\exp(\beta_s[U(x_0)-k])\mu_{\beta_s}(\phi)\,\mathrm{d}s \\
&\leqslant K_5\ln(1+t) + \left|M_t^{\widetilde{\psi}^2}\right| + K_6\widetilde{G}_t(\phi),
\end{aligned}$$

où les K_i, pour $i \geqslant 5$, représentent désormais des constantes (éventuellement aléatoires, mais indépendantes du temps). On aura remarqué que le terme $K_5\ln(1+t)$ n'apparaît que si $x_0 \in \widehat{M} \setminus \widetilde{M}$ (car alors $V(x_0) = k - U(x_0)$), il peut être remplacé par K_5 si $x_0 \in \widetilde{M}$.

Cependant, $M^{\widetilde{\psi}^2}$ est également une martingale à sauts bornés, il existe donc $K_7 > 0$ tel que pour tout $t \geqslant 0$,

$$\left| M_t^{\widetilde{\psi}^2} \right| \leqslant K_7 + 2\sqrt{\left\langle M^{\widetilde{\psi}^2} \right\rangle_t \ln_2\left(\left\langle M^{\widetilde{\psi}^2} \right\rangle_t + e^e \right)}.$$

Or

$$\left\langle M^{\widetilde{\psi}^2} \right\rangle_t \leqslant 4 \left[\max_{0 \leqslant s \leqslant t} \left\| \widetilde{\psi}_s \right\| \right]^2 \left\langle M^{\widetilde{\psi}} \right\rangle_t \leqslant 4K_2^2 \left\langle M^{\widetilde{\psi}} \right\rangle_t$$

et on obtient ainsi que

$$\left\langle M^{\widetilde{\psi}} \right\rangle_t \leqslant K_8 \ln(1+t) + K_9 \widetilde{G}_t(\phi) + K_{10} \sqrt{\left\langle M^{\widetilde{\psi}} \right\rangle_t \ln_2\left(\left\langle M^{\widetilde{\psi}} \right\rangle_t + \mathrm{e}^e \right)},$$

puis

$$\left\langle M^{\widetilde{\psi}} \right\rangle_t \leqslant K_{11} \ln(1+t) + K_{12} \widetilde{G}_t(\phi)$$

et pour tout $t \geqslant 0$,

$$\begin{aligned} \left| M_t^{\widetilde{\psi}} \right| &\leqslant K_3 + \sqrt{\left[K_{11} \ln(1+t) + K_{12} \widetilde{G}_t(\phi) \right] \ln_2 \left[K_{11} \ln(1+t) + K_{12} \widetilde{G}_t(\phi) + \mathrm{e}^e \right]} \\ &\leqslant K_3 + K_{13} \left[\sqrt{\ln(1+t) \ln_2 [\ln(1+t) + \mathrm{e}^e]} \vee \sqrt{\widetilde{G}_t(\phi) \ln_2 [\widetilde{G}_t(\phi) + \mathrm{e}^e]} \right] \\ &\leqslant K_{14} \sqrt{\ln(1+t) \ln_2 [\ln(1+t) + \mathrm{e}^e]} + K_{13} \sqrt{\widetilde{G}_t(\phi) \ln_2 [\widetilde{G}_t(\phi) + \mathrm{e}^e]}. \end{aligned}$$

Notons qu'en temps grand,

$$\widetilde{g}_t(\phi) \gg K_{14} \sqrt{\ln(1+t) \ln_2 [\ln(1+t) + \mathrm{e}^e]}$$

et que

$$\widetilde{g}_t(\phi) + \widetilde{G}_t(\phi) \gg K_{13} \sqrt{\widetilde{G}_t(\phi) \ln_2 [\widetilde{G}_t(\phi) + \mathrm{e}^e]}$$

ce qui permet de conclure que

$$\widetilde{G}_t(\phi) \sim \widetilde{g}_t(\phi) \quad \square$$

Remarquons qu'il serait aussi possible d'améliorer les estimées de vitesses de convergence que l'on peut obtenir pour les points de $\widetilde{M}$ à partir de la démonstration de la section précédente, en utilisant plutôt la généralisation pour les martingales à temps continus de la loi du logarithme itéré dûe à Stout (cf. Stout, 1970).

En fait, les démonstrations précédentes s'adaptent immédiatement pour obtenir les résultats suivants.

Proposition 9.

- *Soit $\bar{\partial}\widehat{M} = \{x \notin \widehat{M} \setminus \exists\, y \in \widehat{M}$ avec $q(x,y) > 0\}$. Si ϕ est l'indicatrice d'un point $x_0 \in \bar{\partial}\widehat{M}$, on a $V(x_0) = 0$ et on obtient que $G_t(\phi)$ reste borné. Cette quantité converge donc en temps grand, car elle est croissante en t.*

- *Soit ϕ l'indicatrice d'un point $x_0 \in \widehat{M}$ et $l < U(x_0) - k$, on note pour $t \geqslant 0$,*

$$G_t^{(l)}(\phi) = \int_0^t \exp(l\beta_s)\phi(X_s)\,\mathrm{d}s$$

Alors comme ci-dessus, $G_t^{(l)}(\phi)$ converge en temps grand.

- *Soit ϕ l'indicatrice d'un point $x_0 \in \widetilde{M}$ et $l \geqslant U(x_0) - k$, outre $G_t^{(l)}(\phi)$ défini comme ci-dessus, on pose pour $t \geqslant 0$,*

$$g_t^{(l)}(\phi) = \int_0^t \exp(l\beta_s)\mu_{\beta_s}(\phi)\,\mathrm{d}s.$$

Alors, $\lim_{t\to+\infty} g_t^{(l)}(\phi) = +\infty$ *et*

$$\lim_{t\to+\infty} (\widetilde{g}_t^{(l)}(\phi))^{-1}\widetilde{G}_t^{(l)}(\phi) = 1.$$

La dernière partie de cette Proposition nous sera utile par la suite. Les deux premières permettent de voir que les points n'appartenant pas à $\widehat{M}$ sont transcients : Soit ϕ la fonction indicatrice de $\widehat{M}$. On a pour tout $t \geqslant 0$,

$$\phi(X_t) = \phi(X_0) + \int_0^t L_s\phi(X_s)\,\mathrm{ds} + M_t^{\phi},$$

où $(M_t^{\phi})_{t\geqslant 0}$ est une martingale dont le crochet oblique $\langle M^{\phi}\rangle$ est donné par

$$\forall\ t \geqslant 0, \qquad \langle M^{\phi}\rangle_t = \int_0^t \Gamma_{\beta_s}(\phi,\phi)(X_s)\,\mathrm{d}s.$$

Notons

$$\overset{\circ}{\partial}\widehat{M} = \{x \in \widehat{M} \mid \exists\ y \notin \widehat{M} \text{ avec } q(x,y) > 0\}$$

alors pour tout $x \in M$,

$$L_s\phi(x) = \phi_{1,s}(x) - \phi_{2,s}(x),$$

avec

$$\phi_{1,s}(x) = 1_{\bar{\partial}\widehat{M}}(x) \sum_{y\in\widehat{M}} q_{\beta_s}(x,y)$$

$$\phi_{2,s}(x) = 1_{\overset{\circ}{\partial}\widehat{M}}(x) \sum_{y\notin\widehat{M}} q_{\beta_s}(x,y)$$

Cependant, d'après le premier point de la Proposition 9,

$$\int_0^t \phi_{1,s}(X_s)\,\mathrm{d}s$$

converge en temps grand, et il en est de même de

$$\int_0^t \phi_{2,s}(X_s)\,\mathrm{d}s$$

d'après le second point (car rappelons que pour $x \neq y$, $q_{\beta_s}(x,y) = \exp(-\beta_s(U(y) - U(x))_+)q(x,y)$ et que pour $x \in \mathring{\partial}\widehat{M}$ et $y \in \bar{\partial}\widehat{M}$, on a $U(y) - U(x) > k - U(x) \geqslant 0$). Ainsi, $\int_0^\infty L_s\phi(X_s)\,\mathrm{d}s$ converge absolument (p.s.). D'une manière similaire, on voit que

$$\langle M^\phi \rangle_\infty = \int_0^\infty \Gamma_{\beta_s}(\phi,\phi)(X_s)\,\mathrm{d}s$$

converge, ce qui entraîne, d'après la Proposition 6, la convergence de M_t^ϕ.

Finalement, on a donc que $\phi(X_t)$ converge en temps grand. La limite est nécessairement (p.s.) 1, car le Théorème 1 montre que les points de $\widetilde{M}$ sont récurrents. En fait, tous les points de $\widehat{M}$ sont récurrents, car d'après la Proposition 8, pour tout $x_0 \in \widehat{M}$,

$$\int_0^\infty \exp(\beta_s[U(x_0) - k])1_{\{x_0\}}(X_s)\,\mathrm{d}s = +\infty.$$

D'où,

Proposition 10. *P.s., les points de $\widehat{M}$ sont récurrents et les autres sont transcients.*

4. Calcul des potentiels dans le cas d'un arbre

On reprend tout d'abord le cas général d'un graphe irréductible et réversible (M, q, μ). On considère l'opérateur L agissant sur $F(M)$ par

$\forall\ \phi \in F(M)$, $\forall\ x \in M$,

$$L\phi(x) = \sum_{y \in M} (\phi(y) - \phi(x))q(x,y)$$

On s'intéresse, pour $\phi \in F(M)$ donné, à l'équation

$$\begin{aligned} \phi(\cdot) - \mu(\phi) &= L\psi(\cdot), \\ \mu(\psi) &= 0. \end{aligned}$$

dont on sait qu'elle admet une unique solution ψ, du fait de l'irréductibilité du graphe.

Modifions un peu les notations. Pour $x, y \in M$, posons

$$\alpha(x,y) = \mu(x)q(x,y) = \alpha(y,x)$$

(par réversibilité du graphe) et

$$\widetilde{\phi}(x) = \mu(x)[\phi(x) - \mu(\phi)]$$

Il s'agit de trouver $\widetilde{\psi} \in F(M)$ telle que pour tout $x \in M$,

$$\sum_{y \neq x} \alpha(x,y)\widetilde{\psi}(y) = \sum_{y \neq x} \alpha(y,x)\widetilde{\psi}(x) + \widetilde{\phi}(x), \tag{9}$$

l'unique solution de (8) étant alors donnée par

$$\psi(\cdot) = \widetilde{\psi}(\cdot) - \mu(\widetilde{\psi}).$$

Cependant, en se basant sur des calculs de Freidlin et Wentzell (voir le Lemme 3.1, p. 177 de Freidlin et Wentzell, 1984), on peut résoudre explicitement des équations du type (9) :

Pour $x \in M$, on appelle arbre en x un ensemble ω de flêches $y \to z$, où $y \neq z \in M$, tel que pour tout $y \neq x$, il existe une seule flêche partant de y (aucune flêche ne part de x) et tel que les flêches ne bouclent pas. On notera Ω_x l'ensemble de ces arbres en x.

Soit $r = (r(y,z))_{y,z \in M}$ une matrice (dont les valeurs diagonales ne joueront aucun rôle). On définit la valuation de ω par rapport à r par

$$V_r(\omega) = \prod_{y \to z} r(z, y),$$

où le produit est pris sur toutes les flêches appartenant à ω, puis la fonction $G_r \in F(M)$ par

$$\forall\, x \in M, \qquad G_r(x) = \sum_{\omega \in \Omega_x} V_r(\omega).$$

L'intérêt de cette fonction est qu'elle est solution de

$$\forall\, x \in M, \qquad \sum_{y \neq x} r(x, y) G_r(y) = \sum_{y \neq x} r(y, x) G_r(x).$$

Cependant, soit $x_1 \neq y_1 \in M$, on définit une nouvelle matrice $\widetilde{r}$ en posant pour tout $(x, y) \in M \times M$,

$$\widetilde{r}(x, y) = \begin{cases} 0 & \text{si } (x, y) = (x_1, y_1), \\ r(x, y) & \text{sinon.} \end{cases}$$

Notons $F^{(r)}_{x_1,y_1}(\cdot) = G_{\widetilde{r}}(\cdot)$ et $f^{(r)}_{x_1,y_1}(\cdot) = r(x_1, y_1) G_r(y_1)[1_{\{x_1\}}(\cdot) - 1_{\{y_1\}}(\cdot)]$, on vérifie facilement à partir de la formule précédente (appliquée avec $\widetilde{r}$ à la place de r), que pour tout $x \in M$,

$$\sum_{y \neq x} r(x, y) F^{(r)}_{x_1,y_1}(y) = \sum_{y \neq x} r(y, x) F^{(r)}_{x_1,y_1}(x) + f^{(r)}_{x_1,y_1}(x).$$

Prenons $r = \alpha$. Par symétrie de cette matrice, il est clair que la fonction G_α est constante. De plus, puisque le graphe (M, α) est irréductible et du fait que tous les éléments de α sont positifs, cette constante (encore notée G_α) est strictement positive.

Soit B un ensemble de couples (x, y) d'éléments de M tels que $x \neq y$ et $\alpha(x, y) > 0$. Pour $b = (x, y) \in B$, notons $f_b(\cdot) = 1_{\{x\}}(\cdot) - 1_{\{y\}}(\cdot)$. Supposons que B est tel que $(f_b)_{b \in B}$ forme une base du sous-espace vectoriel $\{f \in F(M) \setminus \sum_{x \in M} f(x) = 0\}$ (un tel ensemble B existe toujours, par irréductibilité de α, et on a $\mathrm{card}(B) = \mathrm{card}(M) - 1$).

Pour $b = (x, y) \in B$, posons $F_b = F^{(\alpha)}_{x,y}$ et $\alpha(b) = \alpha(x, y)$. Soit $(a_b)_{b \in B}$ les coordonnées de $\widetilde{\phi}$ dans la base $(f_b)_{b \in B}$. Alors, la fonction

$$\widetilde{\psi} = \sum_{b \in B} \frac{a_b}{\alpha(b) G_\alpha} F_b$$

satisfait (9), ce qui termine de résoudre (8).

Supposons maintenant que M est un arbre. Soit $x_0 \in M$ fixé (pour cette section, comme au début de la Section 2, on n'a pas besoin de l'hypothèse $x_0 \in \widetilde{M}$). Introduisons quelques notations :

Pour tout $x \in M$, soit $p_x = (p_x(i))_{1 \leqslant i \leqslant n_x}$ l'unique chemin allant de x à x_0. Si $x \neq x_0$, on note $g_x = p_x(2)$ et on considère les ensembles :

$$E_x = \{y \in M \setminus \forall\ 1 \leqslant i \leqslant n_y,\ p_y(i) \neq x\},$$
$$F_x = M \setminus E_x.$$

On aura également besoin, pour $x \neq x_0$, de la fonction $\chi(x, \cdot\,)$ définie par

$$\forall\ z \in M, \qquad \chi(x,z) = \begin{cases} -\mu(F_x) & \text{si } z \in E_x, \\ \mu(E_x) & \text{si } z \in F_x. \end{cases}$$

Proposition 11. *Soit ϕ l'indicatrice de $\{x_0\}$. L'unique solution de* (8) *est donnée par*

$$\psi(\cdot\,) = \mu(x_0) \sum_{x \neq x_0} \frac{\mu(F_x)}{\alpha(x, g_x)} \chi(x, \cdot\,).$$

Démonstration. Dans le cas d'un arbre "pointé en x_0", il existe un choix canonique pour B, il suffit de prendre

$$B = \{(x, g_x)\,;\, x \neq x_0.\}$$

Soit $b = (x, g_x) \in B$ fixé, il est immédiat de voir que $F_b(\cdot\,) = G_\alpha 1_{E_x}(\cdot\,)$. On en déduit que

$$F_b(\cdot\,) - \mu(F_b) = -G_\alpha \chi(x, \cdot\,).$$

Cependant, d'après les considérations précédentes, on sait que l'unique solution de (8) est donnée par

$$\psi(\cdot\,) = \sum_{b=(x,g_x) \in B} \frac{a_b}{\alpha(x, g_x) G_\alpha} [F_b(\cdot\,) - \mu(F_b)]$$

et on calcule facilement (du fait que l'on a un arbre) que pour $b = (x, g_x) \in B$, $a_b = -\mu(x_0)\mu(F_x)$, d'où le résultat annoncé. $\square$

Soit $s \geqslant 0$. On va en fait appliquer la Proposition 11 avec l'arbre irréductible et réversible $(M, q_{\beta_s}, \mu_{\beta_s})$. On reprend les notations de l'introduction : ψ_s est la solution de (1) (où comme ci-dessus, $x_0 \in M$ est fixé et ϕ est l'indicatrice de ce point). On a:

Proposition 12. *Il existe une constante $K \geqslant 0$, telle que pour tout $s \geqslant 0$,*

$$\|\psi_s\| \leqslant K \exp(\beta_s V(x_0)),$$
$$\|\partial_s \psi_s\| \leqslant \begin{cases} K \exp(\beta_s [V(x_0) - k]) & si\ V(x_0) \neq 0, \\ K \exp(-\beta_s [k + \delta]) & si\ V(x_0) = 0, \end{cases}$$

où

$$V(x_0) = \max_{x \neq x_0} [H(x_0, x) - U(x) - U(x_0)] \geqslant 0.$$

Démonstration. Pour $s \geqslant 0$ et $x, y \in M$, notons

$$\alpha_{\beta_s}(x,y) = \mu_{\beta_s}(x) q_{\beta_s}(x,y) = Z_{\beta_s}^{-1} \exp(-\beta_s[U(x) \vee U(y)])\mu(x)q(x,y)$$

et si $x \neq x_0$,

$$\chi_{\beta_s}(x,y) = \begin{cases} -\mu_{\beta_s}(F_x) & \text{si } y \in E_x, \\ \mu_{\beta_s}(E_x) & \text{si } y \in F_x, \end{cases}$$

$$A_s(x,y) = \mu_{\beta_s}(x_0) \frac{\mu_{\beta_s}(F_x)}{\alpha_{\beta_s}(x, g_x)} \chi_{\beta_s}(x,y)$$

D'après la Proposition 11, on a pour tout $y \in M$,

$$\psi_s(y) = \sum_{x \neq x_0} A_s(x,y)$$

et il suffit donc d'estimer, pour $x \neq x_0$ fixé, $\|A_s(x,\cdot)\|$ (et $\|\partial_s A_s(x,\cdot)\|$).

Cependant, on voit qu'il existe une constante $K(x_0,x) > 0$ telle que quand β_s devient grand,

$$\|A_s(x,\cdot)\| \sim K(x_0,x)\exp(\beta_s V(x_0,x)),$$

où $V(x_0,x) = U(x) \vee U(g_x) - \min_{z \in F_x} U(z) - U(x_0)$.

Posons $V(x_0) = \max_{x \neq x_0} V(x_0,x)$ et $S(x_0) = \{x \neq x_0 \setminus V(x_0,x) = V(x_0)\}$. Alors, en notant $K(x_0) = \sum_{x \in S(x_0)} K(x_0,x)$, on a,

$$\|\psi_s\| \leqslant \sum_{x \neq x_0} \|A_s(x,\cdot)\|$$
$$\sim K(x_0)\exp(\beta_s V(x_0)).$$

Or on vérifie immédiatement que l'on a aussi

$$V(x_0) = \max_{x \neq x_0} e(p_x) - U(x) - U(x_0)$$

d'où la première inégalité annoncée (pour un arbre, rappelons que l'on a $H(x_0,x) = H(x,x_0) = e(p_x)$, puisque $\mathscr{C}_{x,x_0} = \{p_x\}$).

La seconde (pour celle-ci la constante K dépend de x_0 et de k) se montre de la même façon, en tenant compte de la forme particulière de β_s et du fait que $\psi_s(x)$ est une fraction rationnelle (dépendant de x et de x_0) en les variables $(q_{\beta_s}(y,z))_{y,z \in M}$. □

On retrouve ainsi dans le cas particulier d'un arbre le résultat obtenu au début de la Section 2. Mais on peut vérifier ici sur des exemples simples que la constante $V(x_0)$ est en général optimale et qu'il existe $K > 0$ tel qu'en temps grand,

$$\|\psi_s\| \sim K \exp(\beta_s V(x_0)).$$

5. Théorème de la limite centrale

Vu les estimées précédentes, le Théorème 2 n'est pas difficile à obtenir, car il suffit d'utiliser un résultat général sur les martingales dû à Rebolledo (Rebolledo, 1980) et rappelé ci-dessous (voir sa Proposition II.1, on pourra aussi consulter (Jacod et Shiryaev, 1987), Théorèmes 3.11 et 3.12 p. 432):

Proposition 13. *Soit* $(M_s^{(t)})_{0\leqslant s\leqslant 1}$ *une famille indéxée par* $t\in\mathbb{R}_+$ *de martingales (càdlàg.) localement de carré intégrable et issues de 0. Faisons les hypothèses suivantes :*

(a) *Il existe une fonction décroissante* $c:\mathbb{R}_+\mapsto\mathbb{R}_+$ *satisfaisant* $\lim_{t\to+\infty}c(t)=0$, *telle que*

$$\max_{0\leqslant s\leqslant 1}\left|\triangle M_s^{(t)}\right|\leqslant c(t).$$

(b) *Il existe une fonction continue et croissante* $a:[0,1]\mapsto\mathbb{R}_+$ *satisfaisant* $a(0)=0$, *telle que pour tout* $0\leqslant s\leqslant 1$ *fixé,* $\langle M^{(t)}\rangle_s$ *converge en probabilité vers* $a(s)$ *quand* t *devient grand.*

Alors, quand t *devient grand, la loi de* $(M_s^{(t)})_{0\leqslant s\leqslant 1}$ *converge étroitement (pour la topologie de Skorokhod sur l'espace* $\mathbb{D}([0,1],\mathbb{R})$ *des trajectoires càdlàg. de* $[0,1]$ *dans* $\mathbb{R}$*) vers la loi du processus gaussien continu centré de fonction de covariance* $[0,1]^2\ni(u,v)\mapsto a(u\wedge v)$.

Supposons désormais que les conditions du Théorème 2 soient satisfaites, mis à part éventuellement $W(x_0)>0$. Comme précédemment, la fonction ψ_s est l'unique solution de (1) (où ϕ est l'indicatrice d'un point $x_0\in\widetilde{M}$ fixé et $s\geqslant 0$), et $(M_t^{\psi})_{t\geqslant 0}$ est la martingale associée.

A priori, on aurait voulu appliquer la Proposition 13 avec des martingales $M^{(t)}$ de la forme

$$\forall\ 0\leqslant s\leqslant 1,\qquad M_s^{(t)}=\begin{cases}(1+t)^{-[1+k^{-1}\widehat{W}(x_0,k)]\backslash 2}M_{st}^{\psi} & \text{si } \widehat{W}(x_0,k)>-k,\\ t^{-\frac{1}{2}}M_{\exp(st)-1}^{\psi} & \text{si } \widehat{W}(x_0,k)=-k\end{cases}$$

pour une constante $\widehat{W}(x_0,k)$ bien choisie (voir la fin de cette section), mais ces martingales ne vont pas nécessairement satisfaire la condition (a) de cette Proposition. C'est pourquoi d'ailleurs dans le Théorème 2 la loi limite n'est pas toujours une gaussienne centrée non dégénérée.

Pour se ramener à des martingales qui vérifient les deux conditions de la Proposition 13, on décompose pour tout $s\geqslant 0$, la fonction ψ_s en $\psi_{1,s}+\psi_{2,s}$ avec

$$\psi_{1,s}(\cdot)=\sum_{x\in\widetilde{M}\backslash\{x_0\}}A_s(x,\cdot),\qquad \psi_{2,s}(\cdot)=\sum_{x\notin\widetilde{M}}A_s(x,\cdot),$$

où on a repris les notations de la section précédente. Soient M^{ψ_1} et M^{ψ_2} les martingales associées :

$$\forall\ t\geqslant 0,\qquad M^{\psi_{1_t}}=\psi_{1,t}(X_t)-\psi_{1,0}(X_0)-\int_0^t\partial_s\psi_{1,s}(X_s)\,\mathrm{d}s-\int_0^t L_{\beta_s}\psi_{1,s}(X_s)\,\mathrm{d}s$$

et

$$\forall\ t\geqslant 0,\qquad M_t^{\psi_2}=\psi_{2,t}(X_t)-\psi_{2,0}(X_0)-\int_0^t\partial_s\psi_{2,s}(X_s)\,\mathrm{d}s-\int_0^t L_{\beta_s}\psi_{2,s}(X_s)\,\mathrm{d}s.\ (10)$$

On a $M^{\psi} = M^{\psi_1} + M^{\psi_2}$, mais dans un premier temps, on va plutôt s'intéresser à M^{ψ_1}, et plus précisément aux martingales $M^{(1,t)}$ données par

$$\forall\ 0 \leqslant s \leqslant 1, \qquad M_s^{(1,t)} = \begin{cases} (1+t)^{-[1+k^{-1}\widetilde{W}(x_0)]\backslash 2} M_{st}^{\psi_1} & \text{si } \widetilde{W}(x_0) > -k, \\ t^{-\frac{1}{2}} M_{\exp(st)-1}^{\psi_1} & \text{si } \widetilde{W}(x_0) = -k, \end{cases}$$

où rappelons que la quantité $\widetilde{W}(x_0)$ a été définie dans l'introduction par

$$\widetilde{W}(x_0) = \max_{x \neq x_0} [\mathrm{e}^{\widetilde{M}}(p_{x_0,x}) - 2U(x)] - 2U(x_0).$$

Pour voir comment ce nombre apparaît, intéressons-nous au comportement asymptotique de

$$\mu_{\beta_s}(\Gamma_{\beta_s}(\psi_{1,s}, \psi_{1,s}) 1_{\widetilde{M}}).$$

Notons que pour tous $x_1, x_2 \in M \setminus \{x_0\}$ et $s \geqslant 0$,

$$\Gamma_{\beta_s}(A_s(x_1, \cdot), A_s(x_2, \cdot)) \equiv 0$$

(ce qui montre que les martingales associées aux fonctions $s \mapsto A_s(x, \cdot)$ sont orthogonales, pour $x \in M \setminus \{x_0\}$). En effet, soit $y \in M$,

$$\Gamma_{\beta_s}(A_s(x_1, \cdot), A_s(x_2, \cdot))(y) = \sum_{z \in M} (A_s(x_1, z) - A_s(x_1, y))(A_s(x_2, z) - A_s(x_2, y)) q_{\beta_s}(y, z)$$

or, pour tout $x \neq x_0$,

$$A_s(x, z) - A_s(x, y) \neq 0 \iff \{y, z\} = \{x, g_x\},$$

ainsi,

$$\Gamma_{\beta_s}(A_s(x_1), A_s(x_2))(y) \neq 0 \implies \{x_1, g_{x_1}\} = \{x_2, g_{x_2}\}$$

ce qui est impossible à réaliser si $x_1 \neq x_2$.

Par définition de $\psi_{1,s}$, on en déduit alors

$$\Gamma_{\beta_s}(\psi_{1,s}, \psi_{1,s}) = \sum_{x \in \widetilde{M} \setminus \{x_0\}} \Gamma_{\beta_s}(A_s(x, \cdot), A_s(x, \cdot))$$

Cependant, on calcule que pour $x \neq x_0$ et $y \in M$,

$$\Gamma_{\beta_s}(A_s(x), A_s(x))(y) = \begin{cases} \mu_{\beta_s}(x_0)^2 \mu_{\beta_s}(F_x)^2 \alpha_{\beta_s}(x, g_x)^{-1} \mu_{\beta_s}^{-1}(y) & \text{si } y \in \{x, g_x\}, \\ 0 & \text{sinon,} \end{cases}$$

ainsi en notant

$$\widetilde{W}(x_0, x) = U(x) \vee U(g_x) - 2 \min_{z \in F_x} U(z) - 2U(x_0),$$

$$K(x_0, x) = \mu^{-3}(N) \mu^2(x_0) \mu^2(F_x) \alpha^{-1}(x, g_x),$$

avec $\alpha(x, g_x) = \mu(x) q(x, g_x)$ et $N_{F_x} = \{z \in F_x \mid U(z) = \min_{F_x} U\}$, on a en temps grand

$$\Gamma_{\beta_s}(A_s(x), A_s(x))(y) \mu_{\beta_s}(y) \sim K(x_0, x) \exp(\beta_s \widetilde{W}(x_0, x)) 1_{\{x, g_x\}}(y).$$

On vérifie aisément que la quantité $\widetilde{W}(x_0)$ définie précédemment coïncide avec

$$\widetilde{W}(x_0) = \max_{x \in \widetilde{M} \setminus \{x_0\}} \widetilde{W}(x_0, x).$$

En outre, posons

$$T(x_0) = \{x \in \widetilde{M} \setminus \{x_0\} \setminus \widetilde{W}(x_0, x) = \widetilde{W}(x_0)\}$$

et

$$K(x_0) = 2 \sum_{x \in T(x_0)} K(x_0, x)$$

de telle manière que les calculs ci-dessus montrent que

$$\mu_{\beta_s}(\Gamma_{\beta_s}(\psi_{1,s}, \psi_{1,s}) 1_{\widetilde{M}}) \sim K(x_0) \exp(\beta_s \widetilde{W}(x_0)).$$

Faisons l'hypothèse $\widetilde{W}(x_0) \geqslant -k$, qui assure la divergence de $\int_0^\infty \mu_{\beta_s}(\Gamma_{\beta_s}(\psi_{1,s}, \psi_{1,s}) 1_{\widetilde{M}})$ ds ainsi que l'équivalence en temps grand

$$\int_0^t \mu_{\beta_s}(\Gamma_{\beta_s}(\psi_{1,s}, \psi_{1,s}) 1_{\widetilde{M}})\,\mathrm{d}s \sim \begin{cases} \widetilde{\sigma}(x_0, k)(1+t)^{1+k^{-1}\widetilde{W}(x_0)} & \text{si } \widetilde{W}(x_0) > -k, \\ \widetilde{\sigma}(x_0, k) \ln(1+t) & \text{si } \widetilde{W}(x_0) = -k, \end{cases}$$

où par définition $\widetilde{\sigma}(x_0, k) = \begin{cases} (1 + k^{-1}\widetilde{W}(x_0))^{-1} K(x_0) & \text{si } \widetilde{W}(x_0) > -k, \\ K(x_0) & \text{si } \widetilde{W}(x_0) = -k. \end{cases}$

On en déduit, grâce aux Propositions 9 et 10, et du fait que $\widehat{M} = \widetilde{M}$, que

$$\begin{aligned} \langle M^{\psi_1} \rangle_t &= \int_0^t \Gamma_{\beta_s}(\psi_{1,s}, \psi_{1,s})(X_s)\,\mathrm{d}s \\ &\sim \int_0^t \Gamma_{\beta_s}(\psi_{1,s}, \psi_{1,s})(X_s) 1_{\widetilde{M}}(X_s)\,\mathrm{d}s \\ &\sim \int_0^t \mu_{\beta_s}(\Gamma_{\beta_s}(\psi_{1,s}, \psi_{1,s}) 1_{\widetilde{M}})\,\mathrm{d}s \\ &\sim \begin{cases} \widetilde{\sigma}(x_0, k)(1+t)^{1+k^{-1}\widetilde{W}(x_0)} & \text{si } \widetilde{W}(x_0) > -k, \\ \widetilde{\sigma}(x_0, k) \ln(1+t) & \text{si } \widetilde{W}(x_0) = -k, \end{cases} \end{aligned}$$

ce qui montre que l'hypothèse (b) de la Proposition 13 est satisfaite avec pour $0 \leqslant s \leqslant 1$,

$$a(s) = \begin{cases} \widetilde{\sigma}(x_0, k) s^{1+k^{-1}\widetilde{W}(x_0)} & \text{si } \widetilde{W}(x_0) > -k, \\ \widetilde{\sigma}(x_0, k) s & \text{si } \widetilde{W}(x_0) = -k, \end{cases}$$

(par ailleurs, les processus $\langle M^{\psi_1} \rangle$ et M^{ψ_1} convergent en temps grand si $\widetilde{W}(x_0) < -k$).

Il faut maintenant vérifier la condition (a).

Les sauts de la martingale M^{ψ_1} sont donnés par $\triangle M_t^{\psi_1} = \psi_{1,t}(X_t) - \psi_{1,t}(X_{t-})$ et ceux de $M^{(1,t)}$ sont donc bornés par

$$\begin{cases} 2(1+t)^{-[1+k^{-1}\widetilde{W}(x_0)]\setminus 2} \max_{0 \leqslant s \leqslant t} \|\psi_{1,s}\| & \text{si } \widetilde{W}(x_0) > -k, \\ 2t^{-\frac{1}{2}} \max_{0 \leqslant s \leqslant \exp(t)-1} \|\psi_{1,s}\| & \text{si } \widetilde{W}(x_0) = -k. \end{cases}$$

Pour estimer ceci, on procède comme dans la démonstration de la Proposition 12 :

$$\|\psi_{1,s}\| \leqslant \sum_{x\in\widetilde{M}\setminus\{x_0\}} \|A_s(x,\cdot)\|.$$

Cependant, pour tout $x \in \widetilde{M} \setminus \{x_0\}$,

$$\begin{aligned}\|A_s(x,\cdot)\| &= \mu_{\beta_s}(x_0)\mu_{\beta_s}(F_x)\alpha_s^{-1}(x,g_x)\max\{\mu_{\beta_s}(E_x);\mu_{\beta_s}(F_x)\}\\ &\leqslant \mu_{\beta_s}(x_0)\mu_{\beta_s}(F_x)\alpha_s^{-1}(x,g_x)\end{aligned}$$

et on en déduit qu'il existe une constante $K > 0$ telle que pour tout $s \geqslant 0$,

$$\|\psi_{1,s}\| \leqslant K\exp(\beta_s\widetilde{V}(x_0))$$

avec

$$\begin{aligned}\widetilde{V}(x_0) &= \max_{x\in\widetilde{M}\setminus\{x_0\}}[U(x)\vee U(g_x) - \min_{F_x} U - U(x_0)]\\ &= \max_{x\neq x_0}[\mathrm{e}^{\widetilde{M}}(p_{x_0,x}) - U(x)] - U(x_0).\end{aligned}$$

Pour que la condition (a) soit satisfaite, il suffit donc de voir que

$$\widetilde{V}(x_0) < \frac{\widetilde{W}(x_0)+k}{2}$$

ou encore que pour tout $x \neq x_0$,

$$2e^{\widetilde{M}}(p_{x_0,x}) - 2U(x) - 2U(x_0) < e^{\widetilde{M}}(p_{x_0,x}) + k - 2U(x) - 2U(x_0)$$

ce qui est bien vérifié, car par définition de $\widetilde{M}$,

$$\forall\, x \in M\setminus\{x_0\},\qquad e^{\widetilde{M}}(p_{x_0,x}) < k.$$

Reste à s'intéresser à la martingale M^{ψ_2}, et on va pour cela reprendre sa définition donnée par (10).

En utilisant que pour tous $x, y \in \widetilde{M}$ et tout $t \geqslant 0$, on a $\psi_{2,t}(x) = \psi_{2,t}(y)$ (car $x_0 \in \widetilde{M}$ et $z \notin \widetilde{M}$ implique que $\widetilde{M} \subset E_z$, d'où pour tout $z \notin \widetilde{M}$ et tous $x, y \in \widetilde{M}$, $A_s(z,x) = A_s(z,y)$), et le fait que X finit par rester dans $\widetilde{M}$, on voit que l'expression

$$\psi_{2,t}(X_t) - \psi_{2,0}(X_0) - \int_0^t \partial_s\psi_{2,s}(X_s)\,\mathrm{d}s - \int_0^t L_{\beta_s}\psi_{2,s}(X_s)1_{\widetilde{M}^c}(X_s)\,\mathrm{d}s$$

finit par être stationnaire en t. Ainsi, seul le terme

$$-\int_0^t L_{\beta_s}\psi_{2,s}(X_s)1_{\widetilde{M}}(X_s)\,\mathrm{d}s$$

peut éventuellement faire diverger M^{ψ_2}.

Cependant, pour $x \notin \widetilde{M}$ et $y \in \widetilde{M}$,

$$L_{\beta_s}A_s(x,\cdot)(y) = \begin{cases}\mu_{\beta_s}(x_0)\mu_{\beta_s}(F_x)\mu_{\beta_s}^{-1}(y) > 0 & \text{si } y = g_x,\\ 0 & \text{sinon.}\end{cases}$$

Soient

$$\underline{W}(x_0,x) = -\min_{z\in F_x} U(z) - U(x_0),$$
$$\underline{T} = \{x \notin \widetilde{M} \setminus \underline{W}(x_0,x) = \underline{W}(x_0),\}$$

(cet ensemble ne dépend pas de $x_0 \in \widetilde{M}$, par ailleurs rappelons que $\underline{W}(x_0) = \max_{x\notin\widetilde{M}} \underline{W}(x_0,x)$,)

$$\underline{K}(x_0) = \sum_{x\in\underline{T}} \mu^{-2}(N)\mu(x_0)\mu(N_{F_x})$$

de telle manière qu'en temps grand on ait

$$\mu_{\beta_s}(1_{\widetilde{M}} L_{\beta_s}\psi_{2,s}(x,\cdot)) \sim \underline{K}(x_0)\exp(\beta_s \underline{W}(x_0)).$$

La loi des grands nombres (Proposition 9) permet alors d'obtenir :

- Si $\underline{W}(x_0) \geqslant -k$, l'intégrale $\int_0^\infty \mu_{\beta_s}(1_{\widetilde{M}} L_{\beta_s}\psi_{2,s})\,\mathrm{d}s$ diverge et en temps grand

$$\begin{aligned}
&\int_0^t L_{\beta_s}\psi_{2,s}(X_s)1_{\widetilde{M}}(X_s)\mathrm{d}s \\
&\sim \int_0^\infty \mu_{\beta_s}(L_{\beta_s}A_s(x,\cdot)1_{\widetilde{M}})\mathrm{d}s \\
&\sim \begin{cases} \underline{K}(x_0)\ln(1+t) & \text{si } \underline{W}(x_0) = -k, \\ (1+k^{-1}\underline{W}(x_0))^{-1}\underline{K}(x_0)(1+t)^{1+k^{-1}\underline{W}(x_0)} & \text{si } \underline{W}(x_0) > -k. \end{cases}
\end{aligned}$$

- Si $\underline{W}(x_0) < -k$, l'expression

$$\int_0^t L_{\beta_s}\psi_{2,s}(X_s)1_{\widetilde{M}^c}(X_s)\,\mathrm{d}s$$

converge en temps grand.

On en déduit les mêmes résultats pour la martingale M^{ψ_2} :

- Si $\underline{W}(x_0) < -k$, $M_t^{\psi_2}$ converge en temps grand.
- Si $\underline{W}(x_0) = -k$, on a en temps grand

$$M_t^{\psi_2} \sim -\underline{K}(x_0)\ln(1+t)$$

- Si $\underline{W}(x_0) > -k$, on a en temps grand

$$M_t^{\psi_2} \sim -(1+k^{-1}\underline{W}(x_0))^{-1}\underline{K}(x_0)(1+t)^{1+k^{-1}\underline{W}(x_0)}.$$

Supposons que $\underline{W}(x_0) \geqslant -k$ et soit $(M^{(2,t)})_{t\geqslant 0}$ la famille de martingales définies par

$$\forall\ 0\leqslant s\leqslant 1, \qquad M_s^{(2,t)} = \begin{cases} (1+t)^{-[1+k^{-1}\underline{W}(x_0)]}M_{st}^{\psi_2} & \text{si } \underline{W}(x_0) > -k, \\ t^{-\frac{1}{2}}M_{\exp(st)-1}^{\psi_2} & \text{si } \underline{W}(x_0) = -k. \end{cases}$$

Les estimées ci-dessus montrent en fait (en décomposant M^{ψ_2} en deux termes, l'un convergeant et l'autre décroissant, puis en utilisant les Théorèmes 1.14, p. 292 et 2.15,

p. 306 de Jacod et Shiryaev (1987) que le processus $(M^{(2,t)})_{t\geqslant 0}$ converge pour la topologie de Skorokhod sur $\mathbb{D}([0,1],\mathbb{R})$ p.s. (et donc en loi) vers le processus déterministe b donné par

$$\forall\ 0\leqslant s\leqslant 1,\qquad b(s)=\begin{cases}-\underline{m}(x_0,k)s^{1+k^{-1}\underline{W}(x_0)} & \text{si } \underline{W}(x_0)>-k,\\ -\underline{m}(x_0,k)s & \text{si } \underline{W}(x_0)=-k,\end{cases}$$

avec

$$\underline{m}(x_0,k)=\begin{cases}(1+k^{-1}\underline{W}(x_0))^{-1}\underline{K}(x_0) & \text{si } \underline{W}(x_0)>-k,\\ \underline{K}(x_0) & \text{si } \underline{W}(x_0)=-k.\end{cases}$$

Finalement, supposons que $W(x_0)\geqslant 0$ et notons $M^{(t)}$ la famille de martingales définie par

$$\forall\ 0\leqslant s\leqslant 1,\qquad M_s^{(t)}=\begin{cases}(1+t)^{-W(x_0)}M_{st}^{\psi} & \text{si } W(x_0)>0,\\ t^{-\frac{1}{2}}M_{\exp(st)-1}^{\psi} & \text{si } W(x_0)=0.\end{cases}$$

En regroupant les résultats précédents, on obtient que

- Si $1+k^{-1}\widetilde{W}(x_0)>2+2k^{-1}\underline{W}(x_0)$, le processus $-M^{(t)}$ converge étroitement quand t devient grand vers la loi du processus gaussien continu centré de fonction de covariance $[0,1]^2\ni(u,v)\mapsto\widetilde{\sigma}(x_0,k)(u\wedge v)^{W(x_0)}$ (la situation $0=1+k^{-1}\widetilde{W}(x_0)>2+2k^{-1}\underline{W}(x_0)$ n'étant pas possible).
- Si $1+k^{-1}\widetilde{W}(x_0)=2+2k^{-1}\underline{W}(x_0)$, le processus $-M^{(t)}$ converge étroitement quand t devient grand vers la loi du processus gaussien continu de fonction de moyenne $[0,1]\ni u\mapsto\underline{m}(x_0,k)u^{W(x_0)}$ et de fonction de covariance $[0,1]^2\ni(u,v)\mapsto\widetilde{\sigma}(x_0,k)(u\wedge v)^{W(x_0)}$ (la situation $0=1+k^{-1}\widetilde{W}(x_0)=2+2k^{-1}\underline{W}(x_0)$ n'étant pas possible).
- Si $1+k^{-1}\widetilde{W}(x_0)<2+2k^{-1}\underline{W}(x_0)$, le processus $-M^{(t)}$ converge p.s. quand t devient grand vers le processus déterministe

$$[0,1]\ni u\mapsto\begin{cases}\underline{m}(x_0,k)u^{1+k^{-1}\underline{W}(x_0)} & \text{si } \underline{W}(x_0)>-k,\\ \underline{m}(x_0,k)u & \text{si } \underline{W}(x_0)=-k.\end{cases}$$

Il en est de même quand t devient grand pour les processus de fluctuations renormalisées $F^{(t)}$ définis par

$$\begin{aligned}\forall\ 0\leqslant s\leqslant 1,\qquad F_s^{(t)}&=\begin{cases}(1+t)^{-W(x_0)}[G_{st}(\phi)-g_{st}(\phi)] & \text{si } W(x_0)>0,\\ t^{-\frac{1}{2}}[G_{\exp(st)-1}(\phi)-g_{\exp(st)-1}(\phi)] & \text{si } W(x_0)=0,\end{cases}\\ &=\begin{cases}-M_s^{(t)}+(1+t)^{-[1+k^{-1}\widetilde{W}(x_0)]\backslash 2}N_{st}^{\psi} & \text{si } \widetilde{W}(x_0)>-k\\ -M_s^{(t)}+t^{-\frac{1}{2}}N_{\exp(st)-1}^{\psi}\end{cases}\end{aligned}$$

où N^{ψ} désigne le processus constant par morceaux donné par

$$\forall\ t\geqslant 0,\qquad N_t^{\psi}=\psi_t(X_t)-\psi_0(X_0)-\int_0^t\partial_s\psi_s(X_s)\,\mathrm{d}s$$

(dont $t\mapsto G_t(\phi)-g_t(\phi)=\int_0^t L_{\beta_s}\psi_s(X_s)\,\mathrm{d}s$ est la projection prévisible duale).

En effet, avec des notations évidentes, on a $N^{\psi}=N^{\psi_1}+N^{\psi_2}$, or, du fait que le processus de Markov $(X_t)_{t\geqslant 0}$ finit par rester dans $\widetilde{M}$, $N_t^{\psi_2}$ finit par être stationnaire en t, et on vérifie facilement qu'il existe une constante $K > 0$ (déterministe) telle que pour tout $t \geqslant 0$,

$$\left|N_t^{\psi_1}\right| \leqslant K[\exp(\beta_t \widetilde{V}(x_0)) + 1]$$

avec $2\widetilde{V}(x_0) < \widetilde{W}(x_0) + k$. Ainsi, p.s.

$$\begin{cases} \lim_{t\to+\infty}(1+t)^{-[1+k^{-1}\widetilde{W}(x_0)]\backslash 2} \max_{0\leqslant s\leqslant 1}\left|N_{st}^{\psi}\right| = 0 & \text{si } \widetilde{W}(x_0) > -k, \\ \lim_{t\to+\infty} t^{-\frac{1}{2}} \max_{0\leqslant s\leqslant 1}\left|N_{\exp(st)-1}^{\psi}\right| = 0 & \text{si } \widetilde{W}(x_0) = -k, \end{cases}$$

ce qui suffit à montrer, si $W(x_0) > 0$, que pour t grand le processus $((1+t)^{-W(x_0)} N_{st}^{\psi})_{0\leqslant s\leqslant 1}$ (respectivement le processus $(t^{-\frac{1}{2}} N_{\exp(st)-1}^{\psi})_{0\leqslant s\leqslant 1}$, si $W(x_0) = 0$) converge p.s. vers 0.

La convergence en loi des variables $F_t = F_1^{(t)}$ annoncée dans le Théorème 2 découle alors (voir la Proposition 3.14, p. 313 de Jacod et Shiryaev (1987) de la convergence des processus $\widetilde{F}^{(t)}$ et du fait que le processus limite est continu.

Terminons cette section par quelques remarques sur certaines particularités des martingales à sauts qui sont apparues :

Si $\underline{W}(x_0) \geqslant -k$, M^{ψ_2} est un exemple de martingale divergeant p.s. vers $-\infty$. De plus, soient $\underline{M} = \widetilde{M} \sqcup \partial\widetilde{M}$ et pour $x \notin \widetilde{M}$ (et toujours $x_0 \in \widetilde{M}$),

$$e^{\underline{M}}(p_{x_0,x}) = \max_{1\leqslant i\leqslant n_{x_0,x} \setminus p_{x_0,x}(i)\in \underline{M}} U(p_{x_0,x}(i)) \ (= \max_{1\leqslant i\leqslant n_{x_0,x} \setminus p_{x_0,x}(i)\in \partial\widetilde{M}} U(p_{x_0,x}(i))).$$

puis

$$\widehat{W}(x_0) = \max_{x\notin\widetilde{M}}[e^{\underline{M}}(p_{x_0,x}) - 2U(x)] - 2U(x_0).$$

L'hypothèse $\underline{W}(x_0) \geqslant -k$ entraîne que $\widehat{W}(x_0) > -k$ et on calcule facilement qu'il existe une constante $\widehat{K}(x_0) > 0$ telle qu'en temps grand

$$\left\langle M^{\psi_2}\right\rangle_t \sim \widehat{K}(x_0)\exp(\beta_t[k + \widehat{W}(x_0)])$$

ainsi dans ce cas la famille $M^{(2,t)}$ de martingales satisfait p.s. les propriétés suivantes:

$$\lim_{t\to+\infty} [M^{(2,t)}]_1 = 0,$$
$$\lim_{t\to+\infty} \left\langle M^{(2,t)}\right\rangle_1 = +\infty.$$

(En effet, pour la première égalité, on aura remarqué que la martingale M^{ψ_2} est à variation finie et est donc totalement discontinue (cf. Jacod et Shiryaev, 1987, Lemme 4.14, p. 41), ce qui permet de voir que son crochet droit est donné par

$$\forall\ t\geqslant 0, \qquad [M^{\psi_2}]_t = \sum_{0<s\leqslant t} (\triangle M_s^{\psi_2})^2$$

on utilise alors le fait que p.s. M^{ψ_2} n'a qu'un nombre fini de sauts, et pour la seconde égalité, on aura noté que $\widehat{W}(x_0) > k + 2\underline{W}(x_0)$.)

Par ailleurs, soit

$$\begin{aligned}\widehat{W}(x_0,k) &= \widehat{W}(x_0) \vee \widetilde{W}(x_0)\\ &= \max_{x \neq x_0}[e^{\underline{M}}(p_{x_0,x}) - 2U(x)] - 2U(x_0)\end{aligned}$$

on déduit des calculs précédents qu'il existe une constante $\sigma(x_0,k) > 0$ telle qu'en temps grand

$$\begin{cases} \langle M^{\psi}\rangle_t \text{ converge} & \text{si } \widehat{W}(x_0,k) < -k,\\ \langle M^{\psi}\rangle_t \sim \sigma(x_0,k)\ln(1+t) & \text{si } \widehat{W}(x_0,k) = -k,\\ \langle M^{\psi}\rangle_t \sim \sigma(x_0,k)\exp(\beta_t[k+\widehat{W}(x_0,k)]) & \text{si } \widehat{W}(x_0,k) > -k. \end{cases}$$

6. Réversibilité faible de Hajek

Pour prouver la loi des grands nombres pour les temps d'occupation en des points de $\widehat{M}$, on s'est placé dans le cadre le plus simple d'un graphe irréductible et réversible, mais la démonstration peut en fait s'adapter à d'autres situations et notamment aux cas où on fait seulement l'hypothèse de réversiblité faible de Hajek (voir Hajek, 1988): Soit q un noyau irréductible de probabilités de transitions sur M (on note μ l'unique probabilité invariante associée) et U une fonction réelle sur M, on dit que (M,q,μ,U) satisfait la condition de réversibilité faible de Hajek, si l'application H sur $M \times M$ introduite dans la première section ($\mathscr{C}_{x,y}$ et $e(\cdot)$ étant définis de la même manière) est symétrique.

Pour $\beta \geqslant 0$, soit μ_β la probabilité invariante pour le noyau de probabilités de transitions q_β construit comme dans l'introduction.

Il est bien connu (voir Trouvé, 1993, p. 37, Théorème 1.35; Hwang et Sheu, 1992) qu'il existe une fonction $v : M \to \mathbb{R}_+^*$ telle que pour tout $x \in M$, on ait quand β devient grand

$$\mu_\beta(x) \sim v(x)\exp(-\beta U(x)).$$

Cependant, ce résultat, les calculs présentés dans [14] et ceux du début de la section 2 (on aura remarqué que le graphe $(\overline{M},\bar{q})$ et la fonction $\bar{U}$ que l'on y obtient en rajoutant un point à M vérifient encore la condition de réversibilité faible de Hajek) permettent de voir que l'estimation (4) et la Proposition 5 sont encore satisfaites. Le reste des démonstrations est alors identique, car on n'y utilise pas la réversibilité (mais seulement l'invariance) de μ_β par rapport à q_β.

Ainsi, sous la condition de réversibilité faible de Hajek, le Théorème 1 est encore vérifié (notons que cette condition assure aussi que les ensembles $\widetilde{M}$ et $\widehat{M}$ sont fortement connexes).

Par ailleurs, on peut se demander quelles formes devront prendre ces lois des grands nombres dans le cas des algorithmes de recuit simulé généralisés (voir Hwang et Sheu, 1992; Trouvé, 1993; Miclo, 1992), ou si les processus considérés sont à valeurs dans une variété riemannienne compacte et connexe.

Références

O. Catoni, Sharp large deviations estimates for simulated annealing algorithms, Ann. Inst. H. Poincaré 27 (1991) 291–383.

R.L. Dobrushin, Central limit theorem for nonstationary Markov chains I et II, Theory Probab. Appli. 1 (1) et (4) (1956) 65–80 et 329–383.

M.I. Freidlin and et A.D. Wentzell, Random Perturbations of Dynamical Systems (Springer, Berlin, 1984).

N. Gantert, Laws of large numbers for the annealing algorithm, Stochastic Processes Appl. 35 (1990) 309–313.

B. Gidas, Non stationary Markov chains and convergence of the annealing algorithm, J. Statist. Phys. 39 (1985) 73–131.

B. Hajek, Cooling schedules for optimal annealing, Math. of Oper. Res. 13 (1988) 311–329.

C.R. Hwang and S.J. Sheu, Singular perturbed Markov chains and exact behaviors of simulated annealing processes, J. Theoret. Probab. 5 (1992) 223–249.

M. Iosifescu, Finite Markov Processes and their Applications (Wiley, New York, 1980).

M. Iosifescu and R. Theodorescu, Random Processes and Learning (Springer, Berlin, 1969).

J. Jacod and A.N. Shiryaev, Limit Theorems for Stochastic Processes (Springer, Berlin, 1987).

J.G. Kemeny and J.L. Snell, Finite Markov Chains (Van Nostrand, New York, 1960).

D. Lepingle, Sur le comportement asymptotique des martingales locales, in: C. Dellacherie, P.A. Meyer et M. Weil, eds., Séminaire de Probabilités XII, Lecture Notes in Math., Vol. 649 (Springer, Berlin, 1978) pp. 148–161.

P.A. Meyer, Martingales and Stochastic Integrals, Lecture Notes in Math. Vol. 284 (Springer, Berlin, 1972).

L. Miclo, Recuit simulé sans potentiel sur un ensemble fini, in: J. Azema, P.A. Meyer et M. Yor, eds., Séminaire de Probabilités XXVI, Lecture Notes in Math. Vol. 1526 (Springer, Berlin, 1992) pp. 47–60.

L. Miclo, Etude des algorithmes de recuit simulé sous-admissibles, à paraître aux Annales de la Faculté des Sciences de Toulouse (1995).

L. Miclo, Problème de sortie discret et théorèmes limites pour les temps d'occupations du recuit simulé, preprint soumis à Probability Theory and Related Fields (1994).

R. Rebolledo, Central limit theorems for local martingales, Z. Wahrsch. Verw. Geb. 51 (1980) 269–286.

T.A. Sarymsakov, Inhomogeneous Markov chains, Theory Probab. Appl. 6 (1961) 178–185.

W.F. Stout, A Martingale analogue of Kolmogorov's law of the iterated logarithm, Z. Wahrsch. Verw. Geb. 15 (1970) 279–290.

W.F. Stout, Almost Sure Convergence (Academic Press, 1974).

A. Trouvé, Parallélisation massive du recuit simulé, Thèse de doctorat, Université Paris 11, (Orsay, 1993).